



Bioactive Compound Collections: From Design to Target Identification

Laraia, Luca; Robke, Lucas; Waldmann, Herbert

Published in:
Chemical Engineering Progress

Link to article, DOI:
[10.1016/j.chempr.2018.01.012](https://doi.org/10.1016/j.chempr.2018.01.012)

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Laraia, L., Robke, L., & Waldmann, H. (2018). Bioactive Compound Collections: From Design to Target Identification. *Chemical Engineering Progress*, 4(4), 705-730. <https://doi.org/10.1016/j.chempr.2018.01.012>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Review

Bioactive Compound Collections: From Design to Target Identification

Luca Laraia,^{1,2,*} Lucas Robke,^{1,3} and Herbert Waldmann^{1,3,*}

The discovery of bioactive compounds underpins many areas of basic biomedical research and constitutes a large part of medicinal chemistry and chemical biology. Synthetic chemistry is now able to provide almost any drug-like molecule imaginable. Therefore, attention has turned to increasing the biological relevance of the compounds to be used in chemical biology and medicinal chemistry research, as well as maximizing their diversity within this large area of chemical space. In this review, we outline key concepts for the design of biologically relevant compound collections by taking inspiration from nature and natural products. We highlight efficient ways to screen the resulting libraries in order to maximize hit rates and the chance of discovering new modes of action. Finally, we discuss state-of-the-art techniques for the identification of molecular targets of hits identified through phenotypic screening approaches.

INTRODUCTION

The discovery of biologically active molecules to improve human health is among the most important challenges in organic chemistry. Small molecules, which have made up the bulk of therapeutic modalities over the last century, have been supplemented with biologics, including antibodies and nucleic acid derivatives, which show great promise in the clinic. However, small molecules are and will continue to remain excellent tools for the study of human health and disease, as well as potential therapeutics.

In the search for biologically active small molecules, a limiting factor has been the availability of synthetic methods of accessing the necessary complexity required, e.g., in the context of natural products (NPs). However, as a result of advances in synthetic techniques, this challenge has largely been solved. With a broad and well-developed chemical methods toolkit, the focus has increasingly shifted to developing strategies for designing and identifying bioactive compound collections. Key questions in this area are which compounds should be synthesized and what the most efficient ways to make them and assess their bioactivities are.

A good multi-purpose compound library will contain molecules that elicit diverse phenotypes by modulating different targets, i.e., they display a great degree of diversity in bioactivity. Assessing biological diversity *a priori* is a challenging task. By retrospectively analyzing the performance in high-throughput screening (HTS), researchers at Novartis developed a metric termed HTS fingerprint (HTS-FP), which assessed a compound's activity in over 190 HTS campaigns over a period of 10 years.¹ It was found that a compound set with a high degree of diversity in HTS-FP was predictive of further biological diversity in subsequent screens. In addition, compounds with a high degree of similarity in their HTS-FP were also significantly more likely to act via a similar mode of action.² This is particularly crucial for NPs, which are often

The Bigger Picture

The research summarized in this review has implications for two of the UN's sustainable development goals (SDGs). The development of new and improved strategies for obtaining bioactive small molecules will have a significant impact on good health and well-being (UN SDG 3). Because the development of a hit compound to a drug typically falls in the remit of a pharmaceutical company, the research also contributes to industry, innovation, and infrastructure (UN SDG 9). All innovations that increase the efficiency with which the industry can bring a drug to market are highly desirable.

One of the key messages of this review is that compounds can display a great degree of biological diversity without having significantly different structures. As such, working with privileged scaffolds, including natural products, is a fruitful strategy for discovering bioactive molecules. In addition, knowing which assays to use for assessing biological diversity and maximizing the chance of discovering bioactivity is crucial.



poorly covered by chemical descriptors. A key conclusion of their work was that that biological diversity did not necessarily originate from chemical diversity.

In this review, we focus on the design, synthesis, and evaluation of biologically active compound libraries. A variety of strategies for accessing biologically relevant and diverse areas of chemical and biological space are described. Phenotypic screening strategies for maximizing the likelihood of identifying a compound's bioactivity and cell-based techniques for rapid assessment of biological diversity are covered. Finally, state-of-the-art techniques for addressing the inevitable target-identification challenges that are associated with the phenotypic approach are discussed.

NATURAL PRODUCTS: INSPIRATION FOR BIOACTIVE COMPOUNDS

NPs have been, and continue to be, excellent sources of bioactive compounds and clinically used drugs.^{3,4} In addition to using the NPs themselves, simple derivatization reactions can also access analogs with improved potency and biophysical properties. Twenty-eight percent of all currently approved drugs are either NPs or derivatives thereof.⁵ NPs have evolved over time to carry out specific functions for the organisms that produce them.⁶ They are particularly effective agents for defense and signaling purposes, and as such, often regulate key cellular pathways.⁷ In addition, as a result of their biosynthetic pathways and associated feedback mechanisms, NPs and their precursors are able to interact with a range of enzymes. Crucially, NPs are typically not promiscuous, but small variations in their scaffolds and substituents allow them to engage different biomolecular targets with high specificity.^{8,9}

Disadvantages are that NPs are often only present in low amounts from their natural sources, and they may not contain suitable reactive sites for further elaboration and determination of structure-activity relationships. However, advances in synthetic chemistry and total synthesis in particular have enabled the scale-up of NP synthetic routes. A particularly impressive example is the synthesis of eribulin (Halaven), a clinically approved anti-mitotic that is derived from the NP halichondrin B.¹⁰ An interesting deviation from traditional total synthesis is diverted total synthesis. In this approach, the retrosynthetic strategy is planned so that a late intermediate can be diversely functionalized to access a variety of analogs.¹¹

Despite the complexities in their synthesis and derivatization, NPs have had a resurgence in recent years. This is partly due to several analyses that showed that an increase in the fraction of sp^3 carbons (F_{sp^3}) and stereogenic centers correlated well with future success in clinical trials.¹² NPs have a high F_{sp^3} , typically around 0.68, which is significantly higher than the average commercially available compound collection often used in typical large screening campaigns, which has an F_{sp^3} of around 0.37.¹³ Although flat molecules have been developed into valuable drugs, there is a demand for more structural diversity and complexity in the design and assembly of small-molecule libraries.¹⁴ Motivated by the success of NPs, those that are commercially available in large quantities have been derivatized.^{15,16} In addition to simple derivatizations, changes to the core scaffolds of NPs can also lead to increased scaffold diversity and bioactivity. In a recent report, Garcia et al.¹⁷ started from a single NP to obtain a library consisting of 16 different scaffolds. This was achieved by transforming the alkaloid sinomenine by ring distortion reactions (Figure 1). Ring fusion, expansion, cleavage, contraction, (de)aromatization, and rearrangement reactions of the four fused rings of sinomenine generated a library of 65 compounds.

¹Max Planck Institute of Molecular Physiology, Department of Chemical Biology, Otto-Hahn-Strasse 11, 44227 Dortmund, Germany

²Department of Chemistry, Technical University of Denmark, Kemitorvet Building 207, Room 124, 2800 Kongens Lyngby, Denmark

³Technische Universität Dortmund, Fakultät Chemie und Chemische Biologie, Otto-Hahn-Strasse 6, 44227 Dortmund, Germany

*Correspondence: luclar@kemi.dtu.dk (L.L.), herbert.waldmann@mpi-dortmund.mpg.de (H.W.)

<https://doi.org/10.1016/j.chempr.2018.01.012>

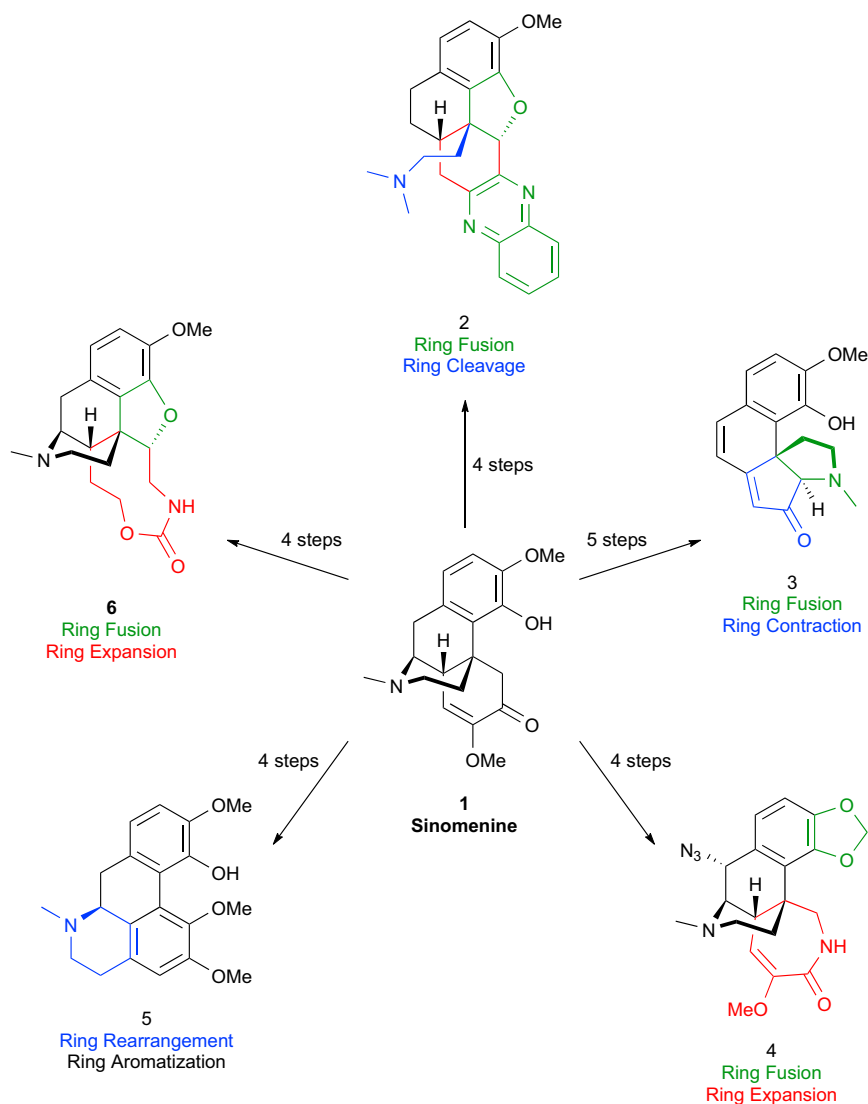


Figure 1. Selected Examples of a Sinomenine-Derived Compound Library

Through ring distortion reactions, the alkaloid NP sinomenine was transformed into diverse scaffolds.¹⁷

Sinomenine (1) is a particularly suitable candidate for the ring distortion and modification of NPs strategy, as it is structurally complex (containing three stereogenic centers and four fused rings) and contains several functional groups that allow further chemical modification. Sinomenine itself has immunosuppressive and anti-inflammatory activities.¹⁸ In this regard, it will be interesting to see what biological activity sinomenine's analogs will present.

The cinchona alkaloids display a broad bioactivity profile, which ranges from the antimalarial quinine to the antiarrhythmic quinidine (Figure 2A, 7 and 8).¹⁹ They are also excellent substrates for the synthesis of compound collections because they contain several orthogonal functional groups for diverse functionalizations. To that effect, a library based on quinine was reported by Hergenrother and co-workers as part of a larger effort to deploy their ring distortion and modification strategy.²⁰ The cinchona alkaloids quinidine (8b) and cinchonine (8a) can be cyclized

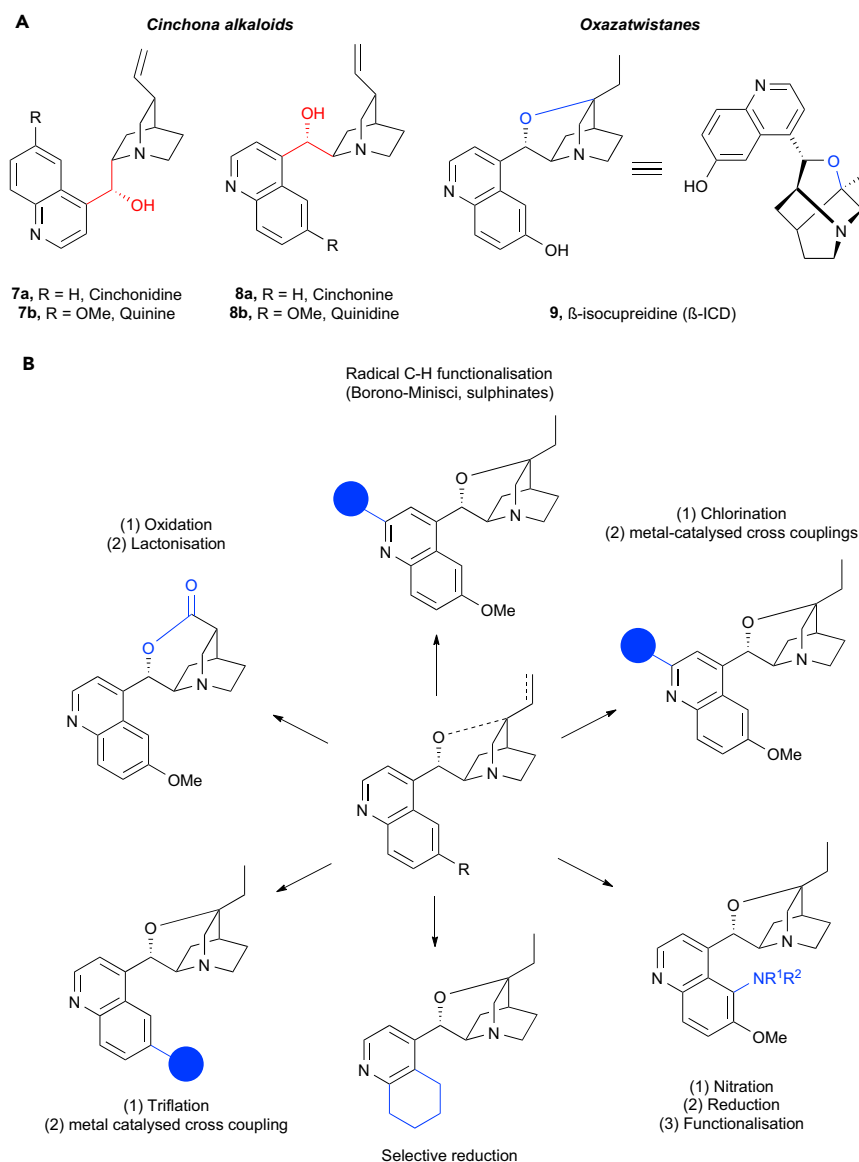


Figure 2. Cinchona-Alkaloid-Derived Compound Libraries

Cinchona alkaloids (A) and the synthesis of an oxazatwistane-derived compound collection (B).

in acidic conditions to yield the highly strained oxazatricyclo[4.4.0.0]decane (oxazatwistane, **9**) scaffold.²¹ It was proposed that the key variation in the scaffold as a result of the cyclization may also endow these compounds with a different bioactivity profile. Therefore, a library of oxazatwistane-derived compounds, with modifications around the quinoline core, was synthesized with C–H functionalization reactions and metal-catalyzed cross couplings (Figure 2B).²²

Overall, a library of 47 compounds was synthesized and subsequently tested in a range of phenotypic assays. This enabled the identification of several potent autophagy inhibitors, which were characterized further to reveal an unusual mode of action. Crucially, none of the parent NPs (**7** and **8**) showed any autophagy inhibitory activity under the screening conditions, and thus the oxazatwistane moiety was essential for activity. Thus, using NPs as starting points for library synthesis and modifying their core scaffolds as well as

their substitution patterns remains a viable strategy for identifying biologically active compounds with interesting modes of action.

BIOLOGY-ORIENTED SYNTHESIS

Although NPs have provided a wealth of drugs, their often challenging synthesis and modification limit their application in modern drug discovery. Biology-oriented synthesis (BIOS) addresses this limitation by guiding the structural simplification of NPs, in order to ease synthesis, while maintaining their biological relevance.^{9,23} BIOS aims to access NP-inspired compound collections without necessarily synthesizing the NP itself. The underlying assumption behind the approach is that if the core scaffold found in the NP is retained in the library, then a substantial portion of the bioactivity should be retained as well. In this context, a scaffold is defined as the combination of ring systems and the linkers that bring them together.⁶ To classify NPs on the basis of their constituent scaffolds, the structural classification of NPs (SCONP) was devised.²⁴ For NPs, the largest scaffold ("child") was deconstructed one ring at a time to deliver the "parent" scaffolds until only one ring system remained. The combination of all scaffolds and their constituent ring systems forms the so-called NP tree (Figure 3). This was subsequently expanded to include non-NPs as well, making it a useful tool for classifying compound collections more generally. The accompanying software to carry out the analysis is freely available at <https://sourceforge.net/projects/scaffoldhunter/files/>.²⁵

The NP scaffold tree can be used to select parent and child scaffolds around which to base compound libraries, which can be particularly efficacious when the accompanying bioactivity data are included. Focusing libraries on the scaffolds not explored previously may cover new intellectual property space in addition to delivering new bioactive compounds.

The BIOS principle was successfully applied to the withanolide class of NPs by Svenda et al.²⁶ The withanolides, a steroid-containing class of NPs comprising more than 300 members, exhibit various bioactivities.^{27–30} A library of small molecules containing a *trans*-hydrindane dehydro- δ -lactone scaffold, which is characteristic of withanolides, was designed and synthesized via the preparation and selective functionalization of three key intermediates (Figure 4A, 12–14). Those key intermediates were then further derivatized by a maximum of two steps leading to final compounds of type 15–17. Subsequent biological investigations identified compounds of type 18 to be inhibitors of the developmental Hedgehog pathway by blocking the function of the smoothened protein (Figure 4B).

Further supporting the versatility of the BIOS principle, another withanolide-inspired and pregnenolone-derived δ -lactone-based compound collection was recently reported to deliver inhibitors of the Wnt pathway (Figure 4B).³¹

FUSION OF NATURAL PRODUCT FRAGMENTS

NPs are promising starting points for the development of compound collections with biological relevance. Therefore, as an extension of the BIOS principle, the fusion of different NP scaffolds offers the possibility to generate compound libraries that consist of multiple biologically pre-validated substructures and should thus lead to an enrichment of potential biological activity.³² A chemoinformatic analysis of 180,000 NPs led to 2,000 fragments, which are diverse, rich in sp^3 -configured

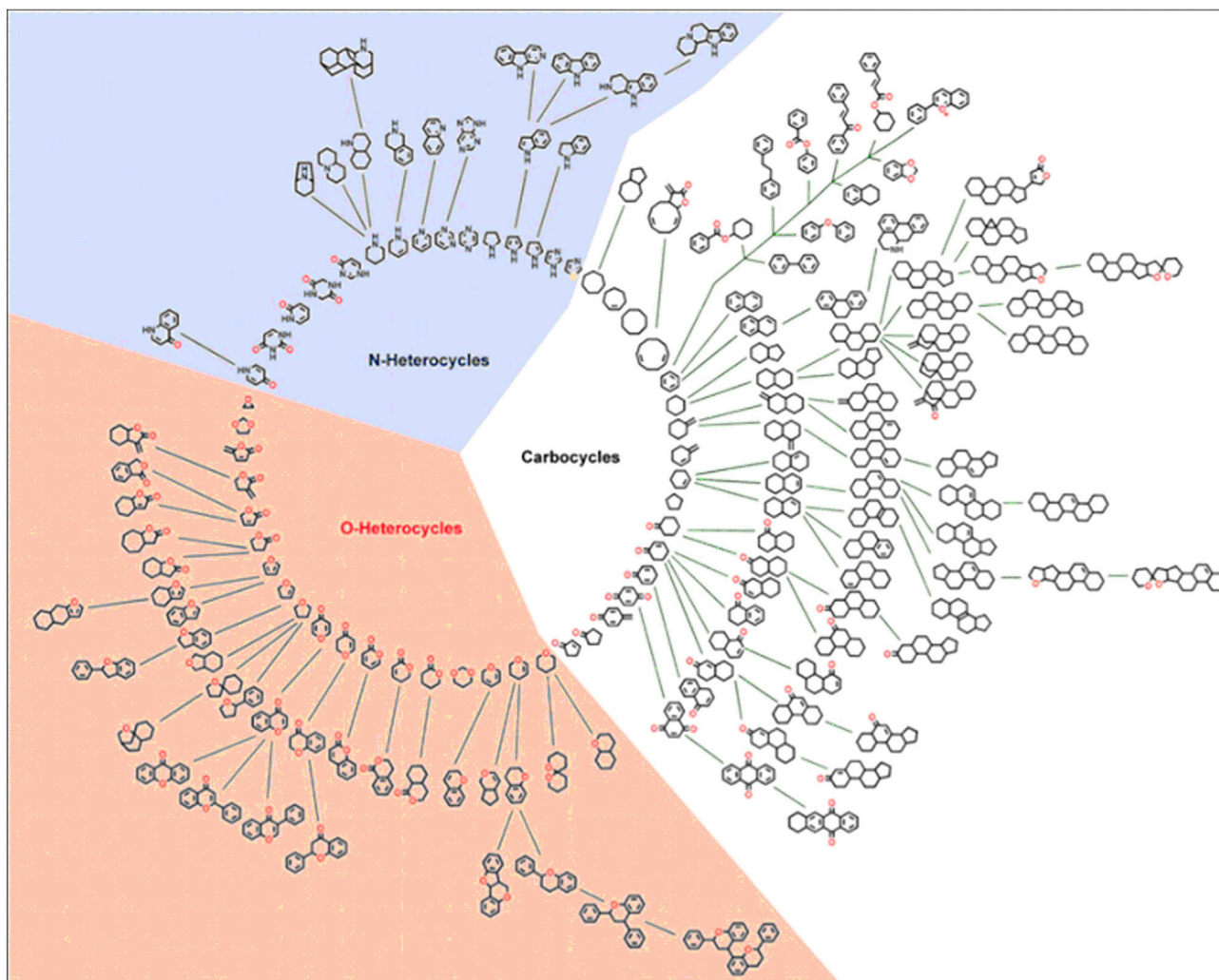


Figure 3. The NP Scaffold Tree

Reproduced from Koch et al.²⁴ Copyright 2005 National Academy of Sciences.

carbons, and similar in their properties to NPs themselves. Their coverage of chemical space is very different from commercially available compound libraries, thus offering more starting points for the identification of novel chemotypes.³³

One such NP fragment is the bicycle tropane. Tropanes can be found in over 600 naturally occurring alkaloids and several of those are used in medical applications, such as the treatment of neurological or psychiatric impairments.^{34,35} However, the complex structure of tropanes poses a challenge for an efficient enantioselective synthesis. A suitable synthetic method has been developed by Narayan et al.³⁶ by means of a Cu^I-catalyzed [3 + 2] cycloaddition of 1,3-fused cyclic azomethine ylides **20** and nitroalkenes **21** (Figure 5A). In this approach, the 1-3-fused ylide reacts as the dipole with the nitroalkene as a comparably low-reactivity dipolarophile to form two new bonds and five stereocenters.

This one-step reaction allowed the generation of a library containing 84 tropane-based compounds. Of the compounds obtained, analogs consisting of an indole-tropane

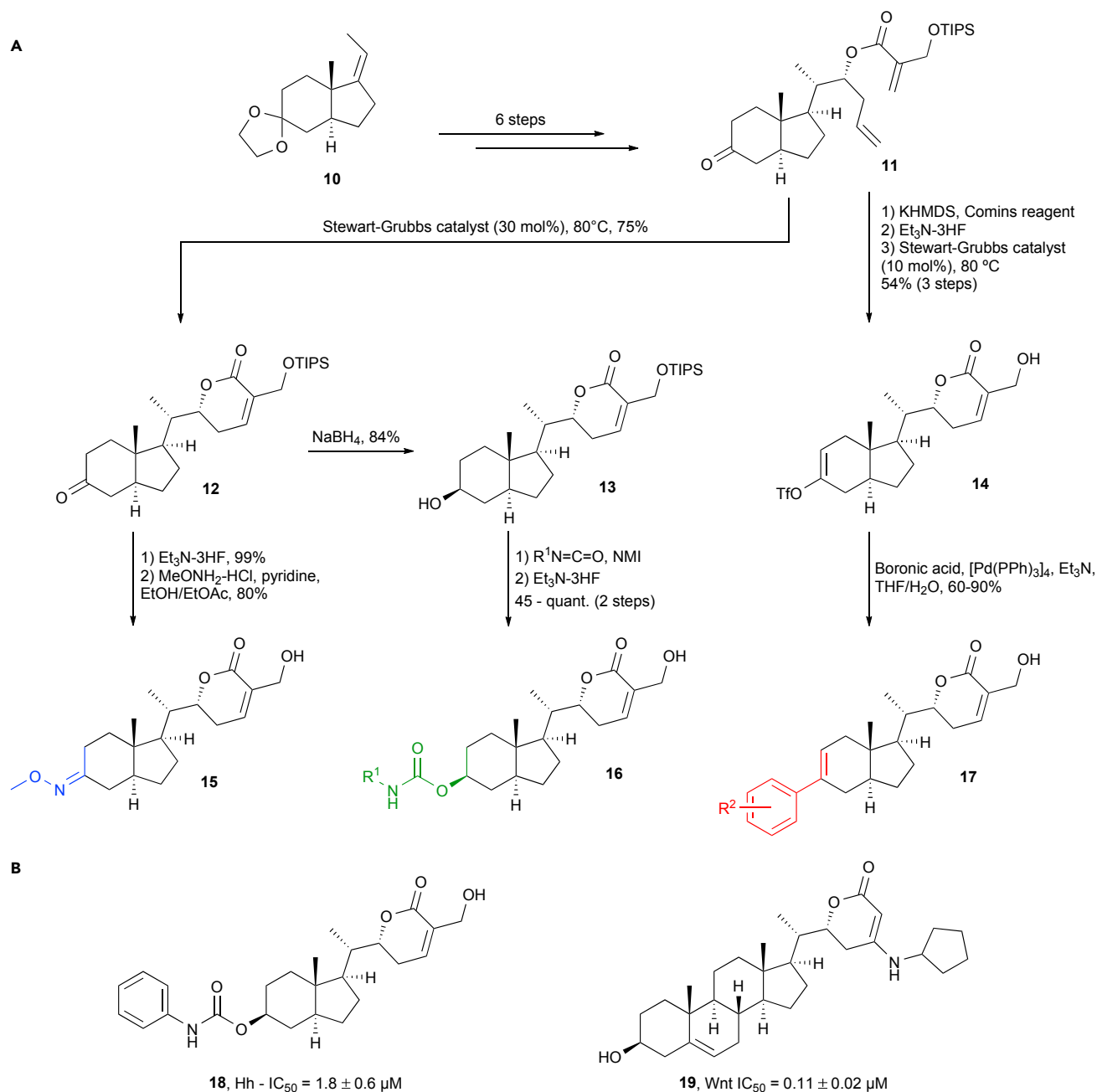


Figure 4. Withanolide-Inspired Compounds

(A) The generation of a withanolide-inspired compound library according to the BIOS principle. The olefin **10** was used as the starting point to yield key intermediates **12–14**, which were transformed into final compounds **15–17**.²⁷

(B) Selected active analogs of the withanolide-inspired compound libraries. Compound **18** was found to be a potent inhibitor of the protein smoothed and thus an inhibitor of Hedgehog signaling. The sterol-derived compound **19** was found to be an inhibitor of Wnt signaling.³¹

scaffold are particularly noteworthy because these NP fusions represent an unprecedented class of hedgehog (Hh) pathway inhibitors (**24** and **25**; Figure 5B).

Another combination of NP fragments was achieved by uniting tropanes and pyrrolidines through a diastereo- and enantioselective cycloaddition reaction. Starting from the racemic tropane **26**, only one enantiomer reacted under dynamic kinetic

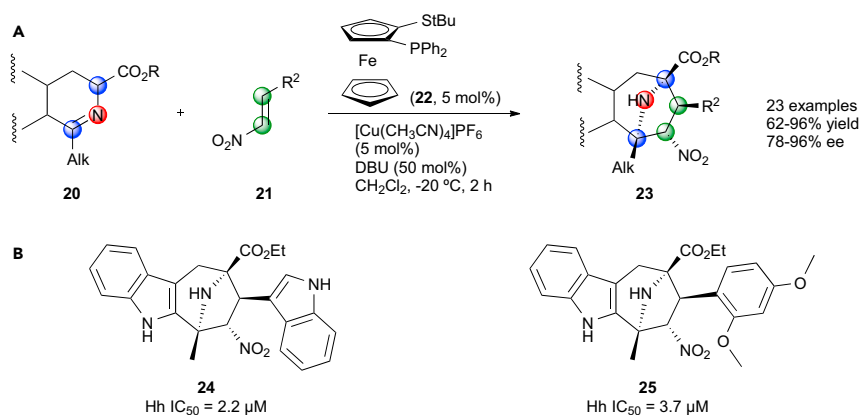


Figure 5. Synthesis and Biological Activity of Indotropanes

(A) Schematic representation of the one-step reaction for the generation of a tropane-based compound library.

(B) Indole-tropane fusion products were found to be novel chemotypes for the inhibition of the Hedgehog (Hh) pathway.

resolution conditions (Figures 6 and 7).³⁷ In a first approach, the 1,3-dipolar cycloaddition of the racemic tropane rac-26 as a dipolarophile with the dipole 27 was established with a copper catalyst and a phosphine ligand 28 (Figure 6). This delivered the desired product 29 as well as the starting material in excellent enantiomeric excess.

Subsequently, enantiodivergent access to both possible absolute configurations of the products 32 and 34 was developed by subjecting the racemic tropane 30 to two different ylides in a sequential manner (Figure 7). In a one-pot reaction, two products with opposing absolute configurations were obtained with only one chiral ligand.

DIVERSITY-ORIENTED SYNTHESIS

A different but related strategy for the design and generation of a diverse library of small molecules, which is not typically based on NPs, is diversity-oriented synthesis (DOS). This strategy has emerged as a powerful approach for generating unbiased screening compounds.³⁸ The key difference to NP-derived and inspired libraries, such as those synthesized by BIOS, is that the key driver is the attainment of compounds that are as chemically diverse as possible. Chemical diversity is typically assessed by cheminformatic analysis, where new libraries are compared with representative sets of NPs (high chemical diversity) and non-NP drugs (low chemical diversity).³⁹ Low indicators of similarity indicate a diverse compound collection. The general hypothesis in this approach is that a large coverage of chemical space will also result in a broader coverage of biological space. Several strategies for accessing diverse compound libraries by DOS have been reported. In the reagent-based approach, a polyfunctional starting material is typically converted to a range of products by simply varying the reaction conditions (Figure 8A).⁴⁰ In the substrate-based approach, a diverse array of substrates is subjected to the same reaction conditions to access diverse molecular frameworks (Figure 8B). To achieve a high degree of chemical diversity, one typically has to access molecules that contain diverse scaffolds. Thus, simply functionalizing one scaffold at different reactive sites would not deliver a library that is characterized by a high degree of chemical diversity. Frequently in combinatorial chemistry, very large libraries consisting of very few scaffolds were prepared, which resulted in low hit rates and low biological relevance. Undoubtedly, DOS holds great promise for bioactive compound discovery, but the

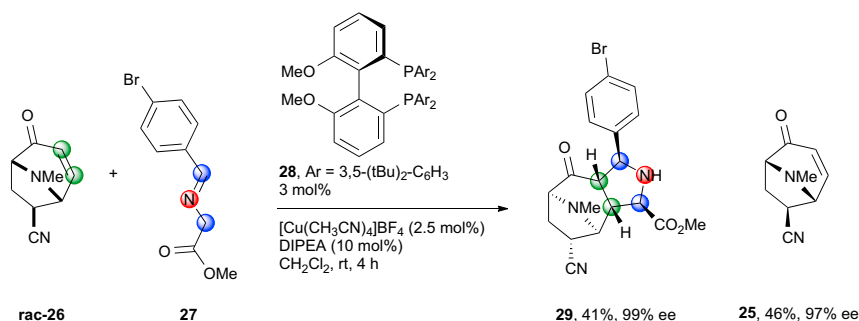


Figure 6. Enantioselective Synthesis of Compounds Embodying Tropanes and Pyrrolidines as NP Fragments

choice and identification of chemical structures that also possess biological activity in a completely unbiased manner remain a challenge.

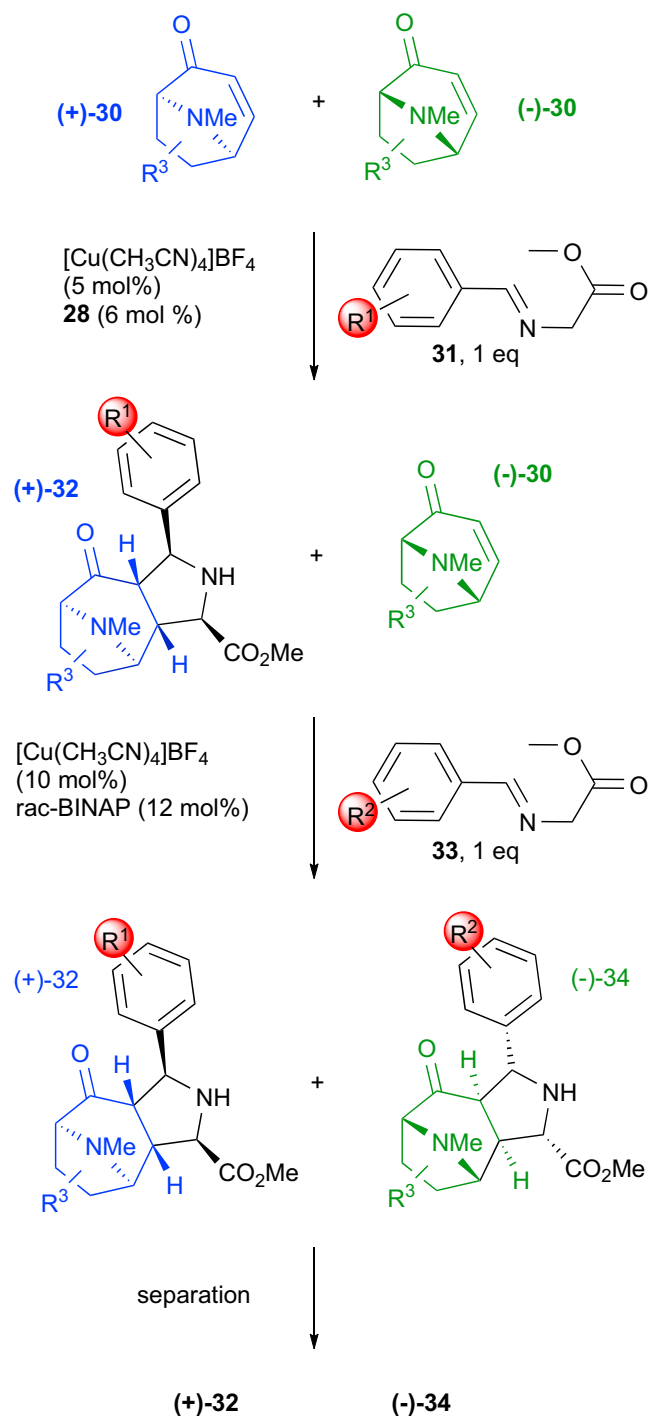
In an attempt to increase the biological diversity and relevance of libraries prepared by DOS, several variations of this approach, such as incorporating NP scaffolds,⁴¹ have been developed. Similarly, another DOS-based approach involves the inclusion of so-called privileged scaffolds or substructures (pDOS) beyond those found in NPs.⁴² In a recent report using a substrate-based pDOS, Kim et al.⁴³ employed a group-pairing strategy of five reactive sites on pyrimidodiazepines **35** (Figure 9). With an average of 2.2 steps, 16 distinct polyheterocycles were prepared, containing pyrimidines and/or diazepines. Both of these are known to possess various biological activities.^{44–46} As a proof of concept, an ELISA-based HTS assay for inhibitors of the protein-protein interaction of leucyl-tRNA synthetase (LRS) and Ras-related GTP-binding protein D (RagD) was used. A compound that specifically inhibited the LRS-RagD interaction was identified, which allowed the regulation of the activation of the mechanistic target of rapamycin (mTORC1).

LIGAND-DIRECTED SYNTHESIS

In most cases, the generation of compound libraries is a laborious endeavor, because the transformation of starting materials is performed in a step-by-step manner in independent experiments, in order to achieve structural diversity. For this reason, new approaches that diversify starting materials in a more unified way would offer great value.^{47–49} In this regard, Lee et al.⁵⁰ developed a synthetic strategy that utilizes different ligands in Au^I catalysis to transform oxindole-based 1,6-enynes to selectively synthesize spirooxindoles, quinolones, or *df*-oxindoles (Figure 10).

The key to this strategy was to identify reactions, which proceed through a common intermediate, whose reaction path can be influenced by different ligands. Therefore, this approach was termed ligand-directed synthesis. Adjusting the ligand's properties, including steric and electronic demand, and fine-tuning the reaction conditions for each transformation could steer the common intermediate's fate down different reaction paths (Figure 11). In the presence of the most electrophilic complex I, cyclopropane ring migration **50** or formation of a benzylic carbocation **51** might be favored, leading to spirooxindole **45**. In contrast, the most bulky ligand, which is also the least electrophilic, is proposed to lead to an oxonium-quinolone gold intermediate **52** by relieving the strain at intermediate **49**. The compromise in steric hindrance and electrophilicity, catalyst II, allows nucleophilic attack by methanol, which opens the cyclopropane ring, eventually yielding compound **47**.

One-pot

**Figure 7. Enantiodivergent Synthesis of (+)-32 and (-)-34**

Only one chiral ligand was used to generate two different enantiomers from racemic tropine 30.

The diverse compound library generated contained compounds that were selective inhibitors of Wnt signaling, Hedgehog signaling, autophagy, or cellular proliferation. Interestingly, these activities depended on the scaffold structure. This finding

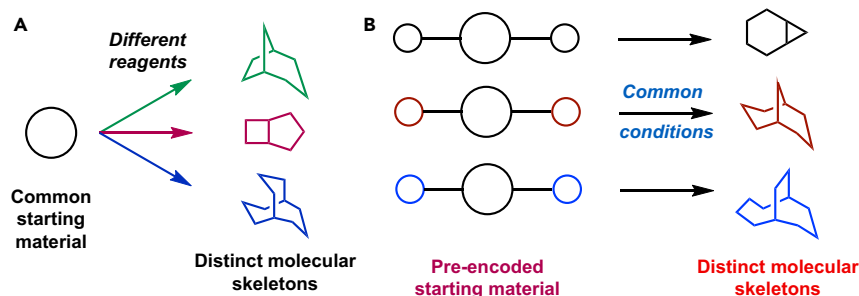


Figure 8. Different Strategies Used in DOS

(A) The reagent-based approach.

(B) The substrate-based approach.

Adapted by permission from Springer Customer Service Centre GmbH: Springer Nature, *Nature Communications*, Galloway et al.,⁴⁰ copyright 2010.

proves the applicability and versatility of ligand-directed synthesis for the generation of compound libraries with privileged scaffolds showing high potential to exhibit biologically active substances.

In addition to gold, other metals have the ability to catalyze diverse processes to obtain various scaffolds from the same starting material, including palladium.⁴⁸ An interesting example of its use to steer reactivity in two opposing directions to yield different products was reported by Alza et al.⁵¹ Starting from the oxidative insertion of palladium into multifunctional intermediate **56**, two possible reaction pathways were observed. Transmetalation with boronic acids led to the direct Suzuki product **59**,⁵² whereas a *syn*-migratory insertion followed by transmetalation led to a 6-5-5-6 fused-ring system (**62**) by a domino Heck-Suzuki-type process (Figure 12).

Varying the palladium source and the ligand enabled the selective synthesis of each scaffold. After screening for anti-mitotic activity, both compounds were found to impair the division of U2OS osteosarcoma cells by different mechanisms.⁵⁴ Whereas the most potent direct Suzuki product (dosabulin, **63**) inhibited tubulin polymerization, the domino Heck-Suzuki product **64** caused chromosome congression defects without affecting tubulin polymerization. It would be interesting to determine the molecular target of this molecule, which is currently unknown.

IDENTIFYING BIOLOGICALLY ACTIVE MOLECULES

Screening a large (>100,000) compound library in a target-based assay typically yields a satisfactory number of hits to pursue. However, focused compound libraries (<1,000 compounds) such as NP-inspired collections, may better be explored by screens that cover broad areas of biological space, where multiple targets are represented, to enable the identification of bioactive compounds. Unbiased phenotypic screening offers this possibility because different potential targets may be associated with a given phenotype.⁵³ In addition, hits identified by phenotypic screening are cell permeable by definition and may act by a polypharmacological mechanism, which may be desirable for particular therapeutic applications, e.g., in certain neurological diseases and cancer.⁵⁴

Image-based screens are particularly valuable because they allow the assessment of cellular morphology and viability in addition to the phenotype of interest. Typical screens involve the use of fluorescently tagged proteins or fluorescent antibodies to a specific protein to image key processes in the cell. Advances in automated

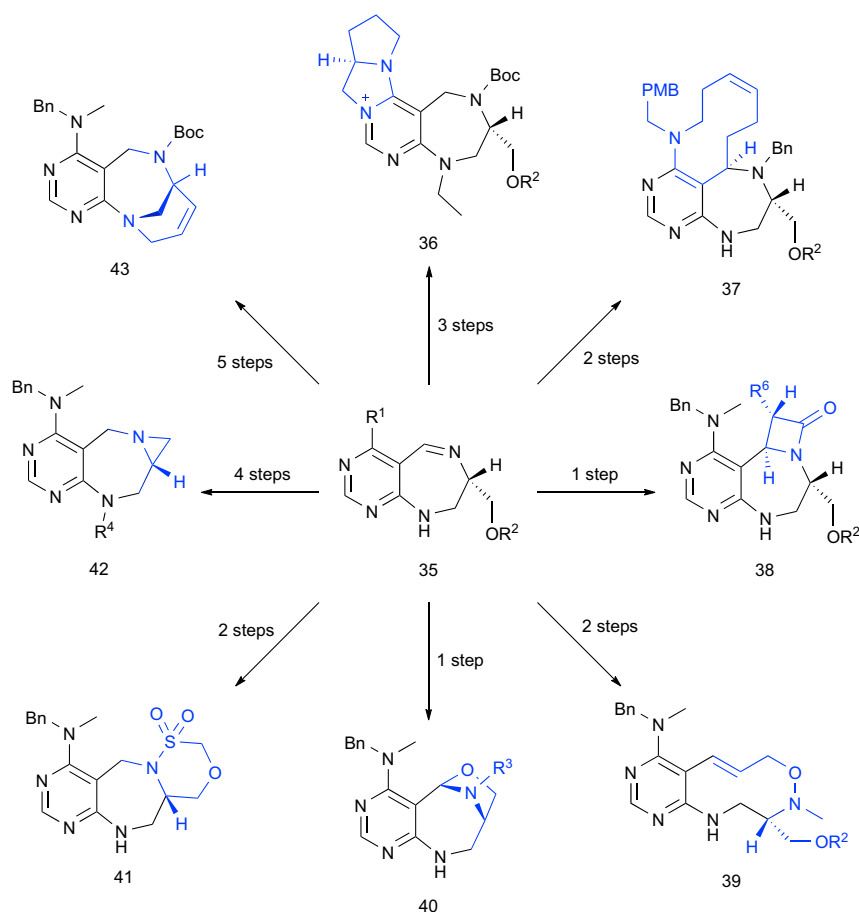


Figure 9. Selected Examples of Pyrimidine- and/or Diazepine-Containing Analogs of a Library Generated by pDOS⁴³

The group-pairing strategy led to compounds containing privileged substructures, one of which inhibited the interaction between LRS and RagD.

image acquisition and analysis have made screening of large libraries in this format practical. Early examples included the identification of new anti-mitotics.⁵⁵ Mitosis can be visualized easily by microscopy because of the large structural rearrangements in a cell during this process. In addition, a wealth of markers for the different mitotic phases exist and have been successfully used. The phosphorylation of histone H3 is a prometaphase marker that is routinely employed to identify cells arrested at this point, after compound treatment. The anti-mitotic activity of dosabulin and HS-1 was identified by this approach (Figure 12).^{51,52}

Another phenotype that can be successfully studied by image-based screening is autophagy. Autophagy is a conserved cellular mechanism for the degradation and recycling of cellular waste material as well as pathogens and is upregulated in situations of stress, such as nutrient deprivation (Figure 13A).⁵⁶ Autophagy inhibition has been proposed as a potential anti-cancer strategy because it may promote tumor survival and growth.⁵⁷ For identifying autophagy inhibitors by phenotypic screening, a cell line stably expressing a GFP-tagged variant of the protein microtubule associated light chain 3 (LC3) can be used, which enables the visualization of autophagosomes in live and fixed cells (Figure 13B). Upon autophagy induction, LC3 localizes to autophagosomes, which can be quantified by image analysis

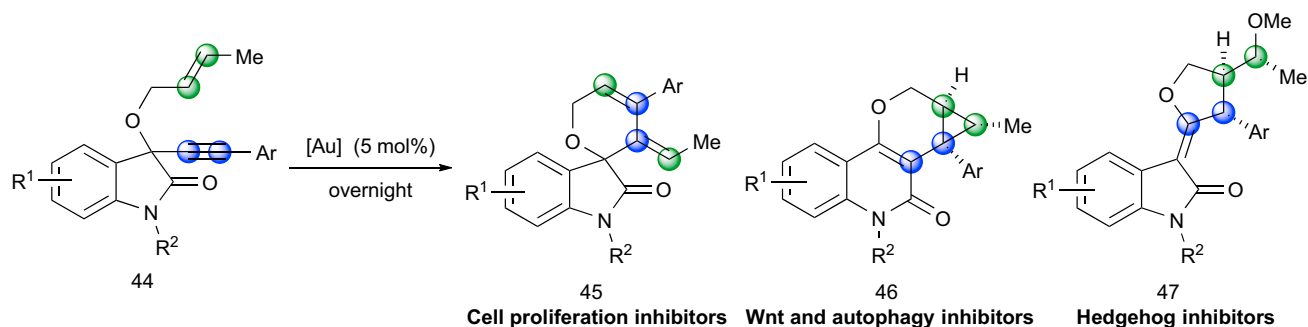


Figure 10. Ligand-Directed Cycloisomerizations of Oxindole-Based 1,6-Enynes 44 to Produce Products Containing Different Scaffolds

(Figure 13B). Recently, a new inhibitor chemotype for the lipid kinase VPS34,⁴⁵ a cinchona alkaloid derivative that acts via a dual mode of action,²² a coumarin with an as yet unknown target,⁵⁸ and a deubiquitinase inhibitor were identified (Figure 13C).⁵⁹

In addition to screening for a specific phenotype, multiparameter imaging or profiling can be used to assess biological activity and diversity in a very general sense and to identify potential targets by similarity to known reference compounds.^{60,61} In “cell painting,” five or six different dyes are used to stain different cellular compartments.⁶² After compound treatment, thousands of parameters are calculated to create a cellular profile or fingerprint for each compound.⁶³ This is compared with the profiles of a large (ideally >5,000) set of reference compounds of known bioactivity and with known targets. Therefore, for every new compound tested, one can determine whether (1) the compound affects the parameters measured and (2) the fingerprint it produces is similar to any reference compound. The comparison with a reference set is particularly useful because it can give an early indication of a mode of action. Even if a compound does not perturb cellular morphology, it may still be biologically active. In fact, a large number of reference compounds show no visible effect in the cell-painting assay. A particularly interesting application of the cell-painting assay has been to determine the biological diversity of a compound collection.⁶⁴ A high degree of diversity in fingerprints in the cell-painting assay was found to be predictive of diversity in future cell-based phenotypic screens, regardless of the phenotype measured. A strategy similar to phenotypic analysis, known as morphobase, which focuses on changes in cellular morphology by using bright-field microscopy images, nuclear morphology, and a nuclear dye, has been reported.⁶⁵

In addition to image-based screening methods, gene-reporter and viability assays are frequently used. For example, gene-reporter assays are used to identify inhibitors of developmental signaling pathways with relevance to cancer, such as Hedgehog (Hh), Wnt, and Notch. The clinically used Hh signaling inhibitor vismodegib was identified by this approach and has become a poster child for the technique.⁶⁶ With this approach, inhibitors of both Wnt and Hh signaling from NP-inspired compound collections were identified (Figures 4B, 5B, and 10).

TARGET IDENTIFICATION

One of the major bottlenecks in phenotypic screening is the identification of the molecular targets of bioactive compounds.⁶⁷ The difficulty of this step depends on a range of variables, which include the novelty of the chemical structure and the cellular localization and abundance of its target. If a compound's structure is

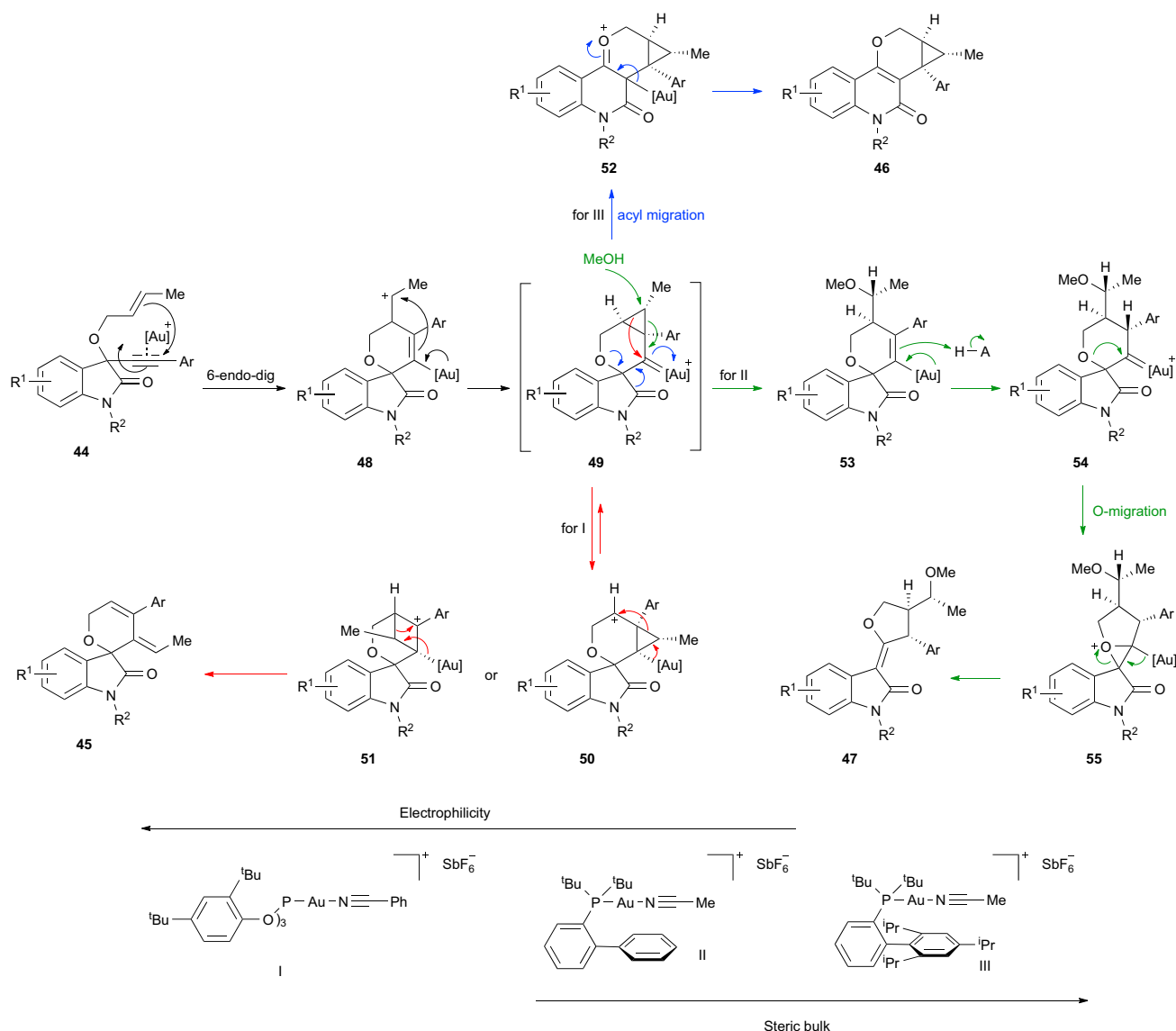


Figure 11. Mechanistic Pathway of the Ligand-Directed Synthesis of Spirooxindoles 45, Quinolones 46, and *df*-Oxindoles 47

The common intermediate was formed by 6-*endo*-dig cyclization and cyclopropane formation starting from O-crotonylated substrate **44**. Depending on the ligand and the reaction conditions, the fate of the intermediate **49** was decided. The most electrophilic Au^I catalyst (**I**) led to spirooxindole **45** (orange reaction path). Quinolones **46** were formed in the presence of the quite bulky gold complex **III** (blue reaction path). The less bulky and less electrophilic gold complex **II** led to the *df*-oxindoles **47** (green reaction path). Adapted from Lee et al.⁵⁰

similar to other bioactive molecules with known targets, this provides a good starting point for target ID. Assessing similarity to known bioactive molecules forms the basis of most computational target prediction tools, including freely available programs (Table 1). These operate by using a range of algorithms to compare the 2D and predicted 3D structures of ligands with those found in databases such as ChEMBL. The similarity ensemble approach (SEA) compares proteins by the set-wise chemical similarity of their ligands.⁶⁸ Swiss target prediction and SuperPred use a combination of 2D and 3D similarity measures to compare a new ligand with known ligands of many thousands of putative targets.^{69,70} ChemProt offers similar target prediction tools, but also enables the user to search by target or even disease type.⁷¹ The self-organizing map-based prediction of drug

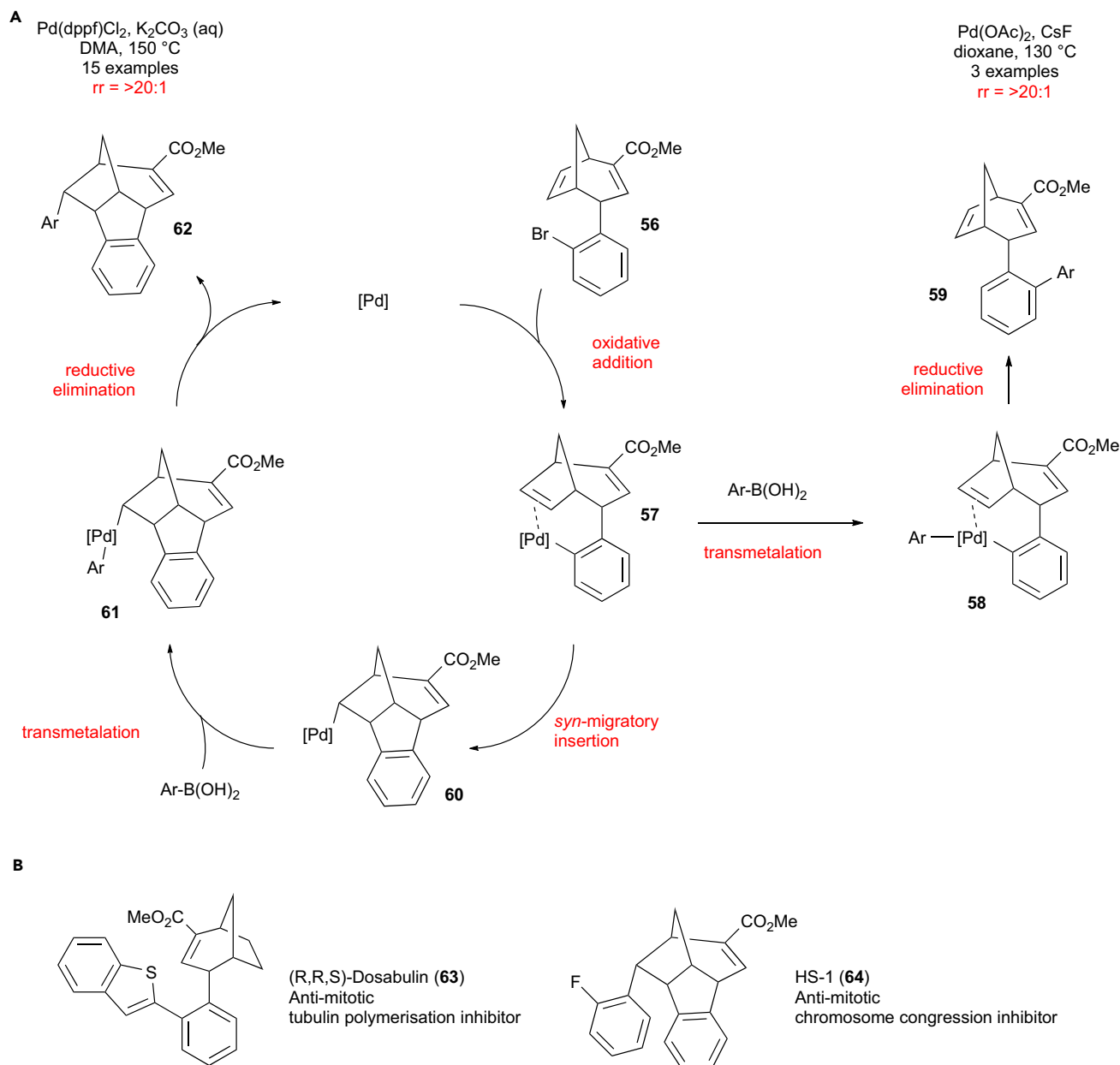


Figure 12. Synthesis and Biological Evaluation of Compounds Derived from Key Polyfunctionalized Intermediate 56 according to a Divergent Pd-Catalyzed Coupling Strategy

(A) Catalytic cycle for the synthesis of the diverse scaffolds.

(B) Selected compounds with anti-mitotic activity.

equivalence relationships (SPiDER) method uses self-organizing maps (SOMs) to relate a query compound to known bioactive substances according to chemical similarity features, such as pharmacophore features and physical and chemical properties.⁷² It may be beneficial to use multiple computational target prediction tools and subsequently select putative targets for validation that are found in several of the approaches used. With SPiDER, it was recently possible to identify the target of the Hedgehog signaling inhibitor Smoothib as the G-protein-coupled receptor (GPCR) Smo.⁷³

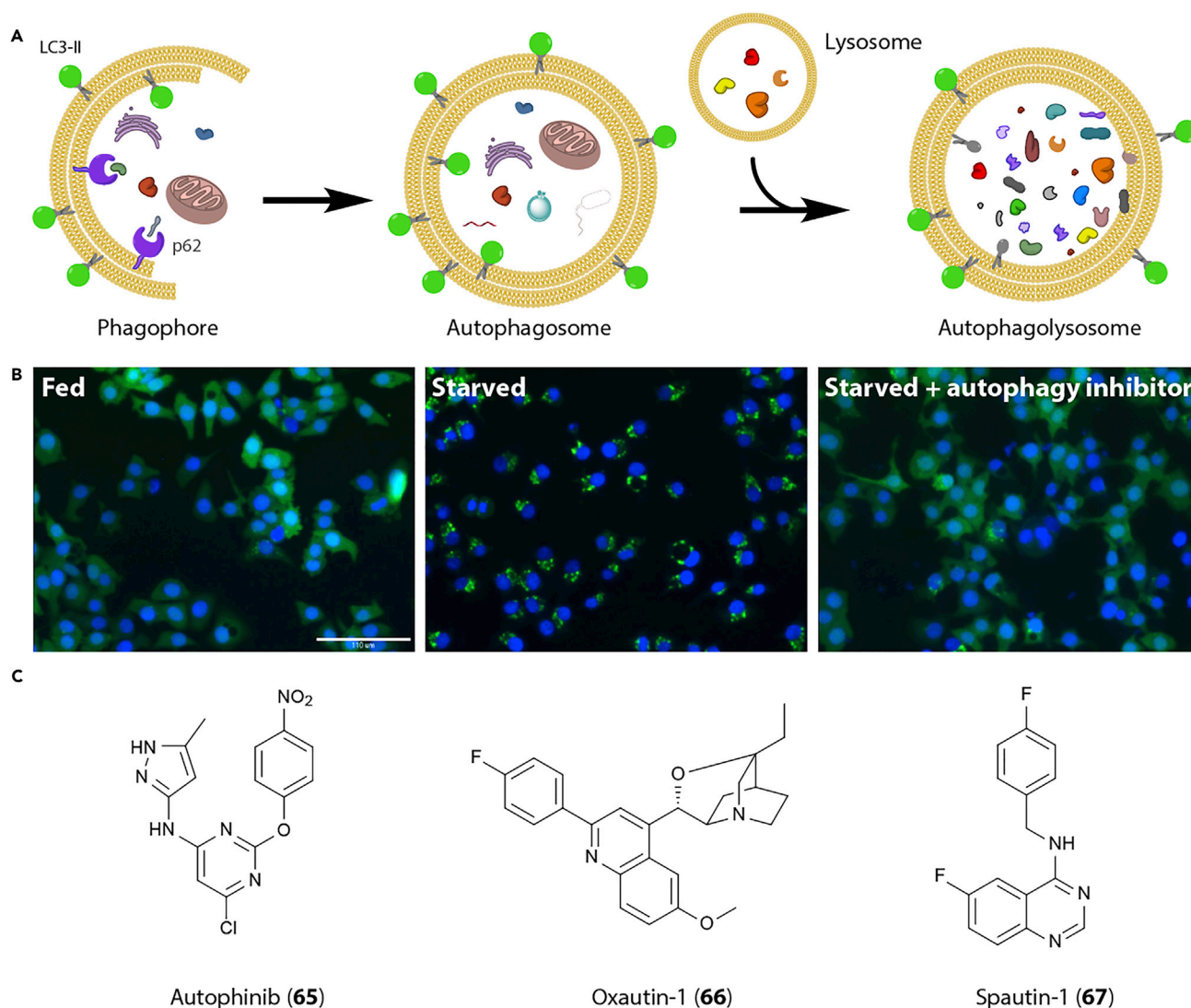


Figure 13. Identification of Autophagy Inhibitors by Phenotypic Screening

(A) Schematic of the autophagic process.

(B) Autophagy was induced by amino acid starvation in MCF7 cells stably transfected with eGFP-LC3. Autophagosomes can be viewed as green puncta or spots, and nuclei are shown in blue. Accumulation of LC3 puncta can be reversed by autophagy inhibitors. Scale bar: 110 μ M.

(C) Autophagy inhibitors identified by phenotypic screening.^{22,37,45}

Computational tools can work very well when the target falls within the limited number that already have a known ligand. These include kinases and GPCRs, which make up a disproportionate number of targets for known drugs. The non-systematic analog of computational target prediction can be described as an educated guess. For example, it may be possible to predict kinase inhibition by visual inspection of the hit molecule, because many kinase inhibitors contain motifs that enable binding to the conserved hinge region of the proteins. For instance, this approach was used to identify the targets of epiblastin, a triamterene-derived pteridine (**68**) that was found to act by inhibiting Casein Kinase 1 (CK1 δ), and a neuritogenic militarone-derived 4-hydroxypyridone (**69**) that was found to target MAP4K4 (Figure 14).^{76,77}

For individual target classes, target prediction may be validated experimentally by means of assays available from commercial suppliers. For instance, various kinase

Table 1. Free and Proprietary Computational Target ID Tools

Software	Producer	Web Address
Similarity Ensemble Approach (SEA)	Kaiser and Shoichet labs ⁶⁸	http://sea.bkslab.org/
Swiss Target Prediction	Michelién lab ⁶⁹	http://www.swisstargetprediction.ch/
ChemProt	Technical University of Denmark and Université Paris Diderot ⁷¹	http://potentia.cbs.dtu.dk/ChemProt/
SuperPred	Structural Bioinformatics Group, Charité ⁷⁰	http://prediction.charite.de/
Target Hunter	Xie lab ⁷⁴	http://www.cbligand.org/TargetHunter/
Chemical Similarity Network Analysis Pull-down (CSNAP)	Torres lab ⁷⁵	http://services.mbi.ucla.edu/CSNAP/
Self-organizing-map-based prediction of drug equivalence relationships (SPiDER)	Schneider lab ⁷²	http://modlabcadd.ethz.ch/software/spider/

panels can be accessed at reasonable cost from companies including Life Technologies and Cerep. In addition, GPCR assays are also available at Cerep and DiscoverRx. However, testing selectivity against a panel of GPCRs is still prohibitively expensive for most labs. Roth and coworkers recently developed a tool that enables high-throughput testing against over 300 non-olfactory GPCRs in parallel. The parallel receptorome expression and screening via transcriptional output, whereby transcriptional activation follows an arrestin translocation (PRESTO-Tango) approach, has been made available from Addgene at an affordable cost (Figure 15).^{78,79}

This approach uses a construct consisting of the GPCR of interest joined to the transcriptional activator (tTA), which contains a TEV protease cleavage site in between. A second construct consisting of the TEV protease linked to human arrestin-2 is transfected simultaneously in a cell line that contains a tTA-dependent reporter gene. Activation of a GPCR results in β -arrestin recruitment, proteolytic cleavage by the linked TEV protease, and freeing of tTA to enter the nucleus and activate reporter genes. Because the association of β -arrestin with ligand-bound receptors is observed for most GPCRs, it was possible to expand this system to a very large portion of all GPCRs. The assay was amenable to a multi-well plate format, and robotics made it possible to screen for agonist activity against almost all non-olfactory GPCRs. The assay format can also be run in antagonist mode for GPCRs with known agonists. The PRESTO-Tango system will facilitate target ID when GPCRs are putative targets and enable testing for compound selectivity across the GPCRome.

When educated guesses and computational target predictions are unsuccessful, the go-to methods are typically mass spectrometry (MS)-based proteomic approaches. The corresponding typical affinity-isolation experiment ("pull-down") involves the synthesis of a suitably functionalized derivative of a hit compound, which can be immobilized onto solid phase and incubated with cell lysates to capture bound proteins non-covalently. Advances in mass spectrometers and analysis tools have made such methods very sensitive, and the use of stable isotope labeling with amino acids in cell culture (SILAC) enables quantification of individual peptide signals and increases the confidence in the hits obtained.⁸⁰ Progress in these and similar proteomic experiments has been reviewed extensively and is thus not covered here.^{67,81,82} A major drawback of the typical pull-down experiment is the

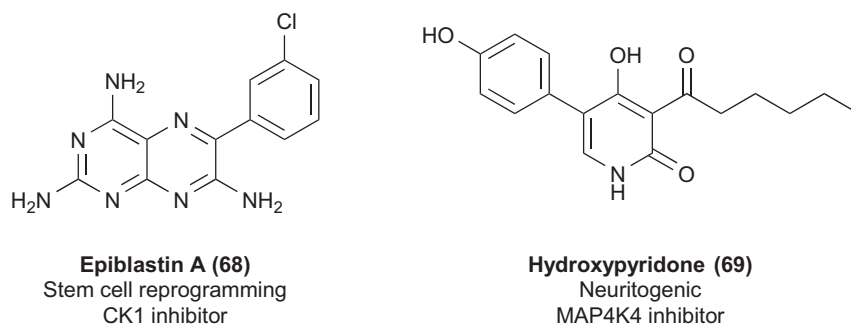


Figure 14. Kinase Targets of Selected Compounds Identified by Educated Guesses

See Ursu et al.⁷⁶ and Schröder et al.⁷⁷

requirement of using cell lysates rather than intact cells, which leads to the exclusion of many, if not all, membrane proteins, including the aforementioned GPCRs. To circumvent this issue, bifunctional probes containing a photoreactive functional group, such as a diazirine, an aryl azide, or a benzophenone, as well as a bio-orthogonal handle for further functionalization to aid subsequent detection or isolation, can be used (Figure 16).⁸³

The addition of small functional groups such as diazirines or alkynes is often tolerated biologically and enables the probe to be administered to live cells, where it can bind targets in their native environment, including membranes (Figure 17). After irradiation to release the reactive intermediate, the probe binds its target covalently. Subsequently, it can be coupled via, e.g., a copper or strain-promoted click reaction to a linker coupled to a fluorophore for in-gel visualization or a biotin group for isolation. As the target is covalently bound, more stringent washing to remove proteins that bind non-specifically to the solid phase is possible. This approach is also amenable to SILAC, providing a robust setup for identifying a wide range of potential targets. In addition to a direct MS readout, a method termed fluorescence difference in two-dimensional gel electrophoresis (FITGE) uses differential labeling of positive and negative bifunctional probes with Cy3 and Cy5 dyes.⁸⁴ A 2D gel separates all proteins by size and pH, and those that are enriched in the positive bifunctional probe can be excised and analyzed.

A recent application of this technique by the Cravatt lab has pushed this technology to the limit of what is currently feasible.⁸⁵ By coupling a bifunctional handle containing a diazirine and an alkyne to a small fragment collection, they were able to identify target proteins of 14 basic fragments in an unbiased cell-based profiling. By increasing the number of fragments to 465 and slightly increasing their complexity and molecular weight (to an average of 340 Da), a phenotypic screening set was established. This contained matched pairs of probes with and without the bifunctional handle. From a subset of the elaborated fragment library containing 300 compounds, three were able to induce differentiation of preadipocytes to adipocytes at 50 μ M, including compound 70 (Figure 18). With the corresponding bifunctional probe 71 and chemical proteomics, the target was identified as progesterone receptor membrane component 2 (PGRMC2). Compound 70 was subsequently characterized as a positive regulator of PGRMC2, and the role of this protein in adipogenesis was validated further by genetic methods.

One of the disadvantages of the proteomic strategies outlined is the requirement for the synthesis of a suitably functionalized pull-down probe. To circumvent this

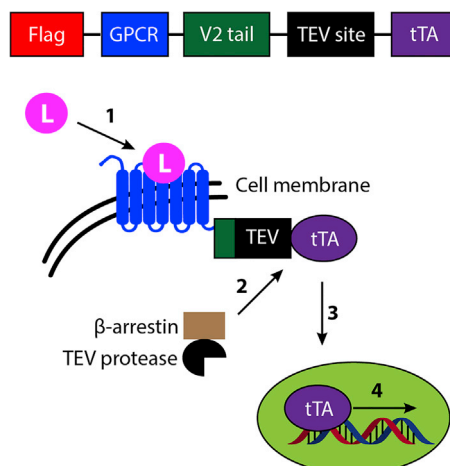


Figure 15. The PRESTO-Tango GPCR Profiling Platform

The design of the construct is shown on top, and the assay format is depicted below. Upon ligand binding (1), the β -arrestin-TEV protease fusion protein is recruited to the GPCR (2) and cleaves off the tTA group. This is then free to translocate to the nucleus (3) and activate the tTA reporter gene (4), which eventually gives rise to a luminescent signal.⁷⁸

problem, thermal proteome profiling (TPP) was developed (Figure 19).⁸⁶ Originally intended as a target validation strategy for showing target engagement in cell lysates, it is now increasingly being used in the target ID process.⁸⁷ Cell lysates are heated to different temperatures in the presence or the absence of a hit compound, and after digestion and labeling with up to 11 isotopically different labels, known as tandem mass tags (TMTs), samples are recombined and analyzed by MS. This enables the determination of melting curves across the whole proteome (mel-tome). Proteins that are significantly stabilized or destabilized by a hit compound are classed as putative targets. TMT labeling is becoming more common for a range of proteomic applications beyond target identification because of the ability to multiplex a large amount of different samples or treatment conditions.⁸⁸ Similar results can also be achieved with the isobaric tag for relative and absolute quantification (i-TRAQ). Disadvantages of the TPP approach include the fact that the stability of some proteins is not significantly altered by compound binding and that this type of experiment typically yields a large list of putative hits, increasing the challenges of target validation.

A variation of this approach that does not require TMT labels is the thermal-stability-shift-based fluorescence difference in two-dimensional gel electrophoresis (TS-FITGE).⁸⁹ In this approach, DMSO- or compound-treated samples heated to the same temperature are centrifuged, differentially labeled with Cy3 and Cy5 dyes, and recombined. A 2D gel separating proteins according to size and charge is run for each temperature, and differences in fluorescence intensity are quantified. In this way, a melting curve for each spot can be generated, and those that show differences between compound treatment and DMSO can be excised and analyzed by MS. This approach identified the targets of the complex NP bryostatin.

A target ID strategy that works by measuring changes in protein levels on a proteome-wide scale is chemproteobase.⁹⁰ In this approach, cells are treated with a hit compound in parallel to a DMSO control. After cell lysis, samples are labeled with two different fluorophores and then recombined. Proteins are separated on a 2D gel, enabling the visualization of changes in fluorescence between the DMSO control

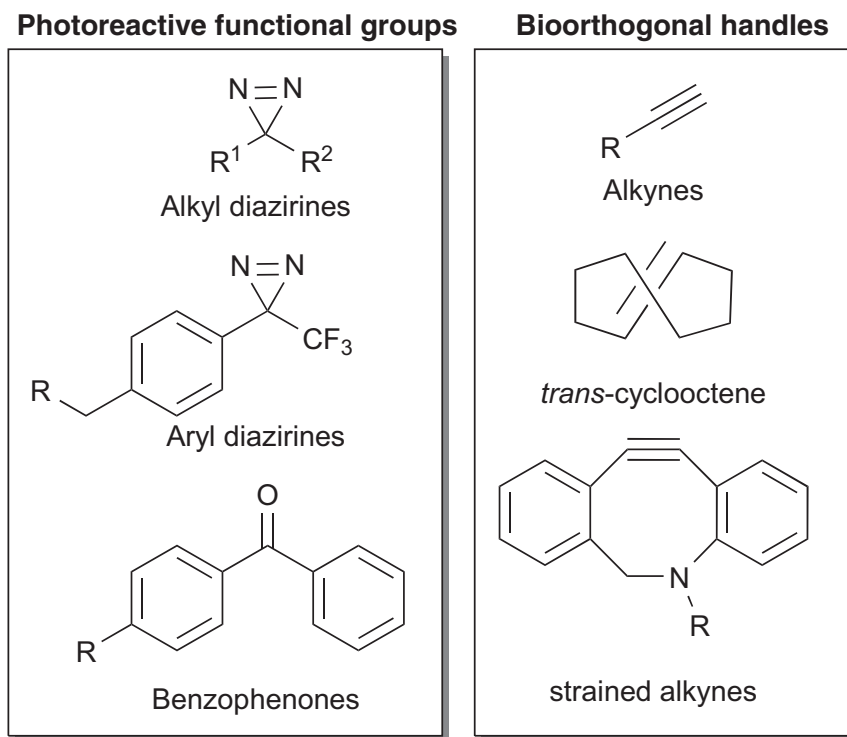


Figure 16. Photoreactive Functional Groups and Bioorthogonal Handles for Isolation or Labeling

and the compound-treated samples. Automated image analysis is used to compare the profile of cells treated with a hit compound with a reference set of compounds with known bioactivities. This allows the user to predict whether the compound acts via a similar mechanism of action to a reference compound. This approach was used to identify the target of the NP pyrrolizilactone as the proteasome and the target of the autophagy inhibitor autophinib as the lipid kinase VPS34.^{45,91}

In addition to chemical, computational, and proteomic methods for target identification, genetic methods have been used as potential strategies for identifying targets for bioactive small molecules. Several strategies have relied on yeast as a simplified model organism for target ID. Haploinsufficiency profiling (HIP) relies on the deletion of one allele of each gene in the yeast genome. For the target of a small molecule, this should result in increased sensitivity to the compound. Homozygous profiling (HOP) involves the deletion of both alleles of a given gene. Although this does not directly identify a target, it can provide information on the network of interacting genes. Both approaches (HIP and HOP) are often combined for increased information on putative targets and their network connectivity. As these strategies have been the focus of extensive reviews, they are not discussed further in this work.⁹² One key disadvantage of yeast-based target ID approaches is that yeast genetics can differ significantly from human genetics, and as such, one could miss a potential target when profiling compounds in yeast. Therefore, in recent years, an increased focus has been placed on identifying new genetic approaches for target ID that can be carried out in mammalian cells.⁹³

An important step forward was the discovery of a near-haploid chronic myelogenous leukemia cell line termed KBM-7.⁹⁴ This has given rise to haploid genetic screening platforms that can be used for target ID methods.^{95,96} By heavily mutagenizing these

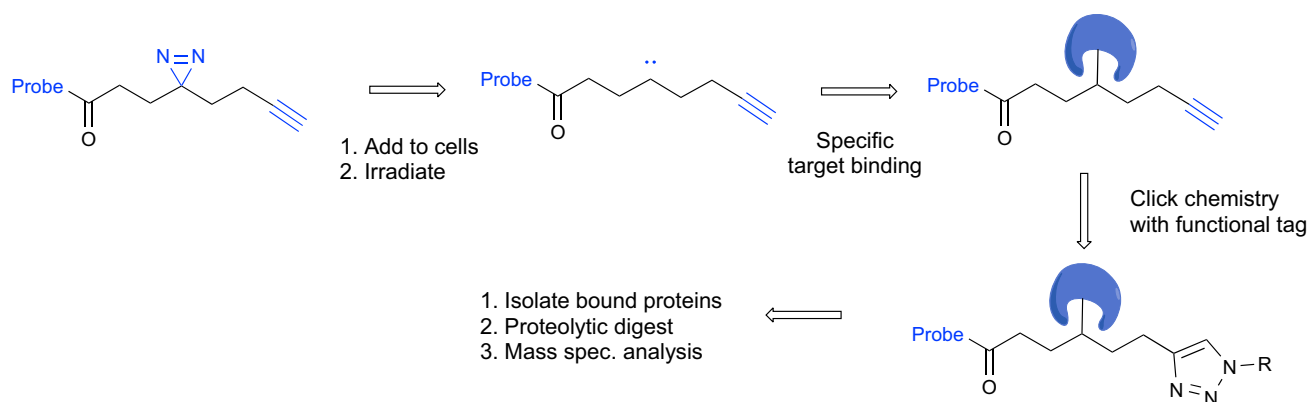


Figure 17. General Strategy for a Photo-affinity Labeling Pull-Down

A bifunctional probe is incubated with cells and irradiated to generate the reactive carbene. This reacts specifically with proximal target proteins. After click chemistry with a functional handle, e.g., biotin, bound proteins can be enriched on a streptavidin-coated solid phase. After proteolytic digest on the retained proteins, these can be analyzed by their constituent peptides via MS.

cells, mutations in every potential gene are present in a large population. Compounds with unknown modes of action that are also cytotoxic are tested in this cell line at a lethal concentration. Colonies that are able to grow in these conditions are sequenced and analyzed for mutations. Mutations can occur in the target protein; however, they may also occur in transfer proteins necessary for the uptake of the molecules or metabolic enzymes that render the small molecule inactive.⁹⁷ A limitation of this technique is that it requires a toxic molecule for resistant colonies to be generated; however, if this prerequisite is satisfied, it is a valuable addition to the target ID toolkit.

A similar approach for the target ID of cytotoxic compounds is the use of the Cancer Cell Line Encyclopedia (CCLE).⁹⁸ This resource contains 947 human cancer cell lines, which have been subjected to gene expression profiling, chromosomal copy number analysis, and massively parallel sequencing. This was initially assembled to predict which tumor types were most likely to respond to specific treatments. However, the strategy was recently adapted for target ID purposes. The NP englerin A was tested in 524 cell lines from the CCLE and was found to only be toxic in a specific subset.⁹⁹ This correlated strongly with mRNA levels of the short transient receptor potential channel 4 (TRPC4). As siRNA knockdown of TRPC4 did not correlate with compound activity, it was speculated that the compound may act as an agonist rather than an antagonist of this calcium channel. This target hypothesis was independently validated by two research groups.^{99,100}

OUTLOOK

Although great advances have been made in the design and synthesis of bioactive compound collections, as well as their biological evaluation, significant challenges still remain, which are currently being addressed to different degrees. Given the vast amount of possibilities, the selection of scaffolds on which to base compound collections remains a key discussion point and challenge. Interestingly, since the introduction of DOS, BIOS, and variations thereof, very few new principles for the design of compound libraries enriched in bioactivity have been reported. New concepts, including the fusion of NP fragments to create new scaffolds or pseudo NPs, are beginning to gain traction. These efforts will benefit greatly from increases in technologies for automating the synthesis process. Automated systems that can

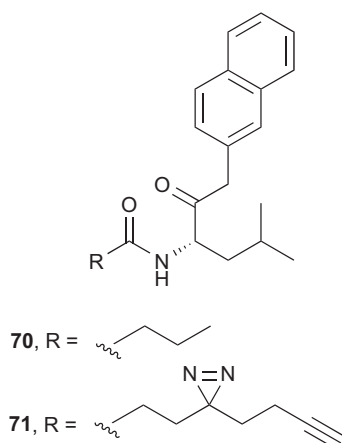


Figure 18. A Tool Compound for PGRMC2 Using Fragment Screening in Cells Followed by Target ID Using Chemical Proteomics with a Matched Bifunctional Probe

See Parker et al.⁸⁵

cater for a large array of reaction types will ensure that more time is made available for assessing a compound's bioactivity.

An additional area of debate in the field is how to select compounds to maximize diversity in a compound library. In this respect, computational predictions of chemical diversity are not necessarily predictive of biological diversity. However, predicting biological diversity *a priori* still remains a major challenge. The introduction of the "cell-painting assay" has delivered a powerful tool for assessing a compound's biological profile, provided that it causes a change in cellular morphology for the organelles stained. Diverse fingerprints in this assay can be used as a metric for the construction of focused chemical libraries that cover a broad range of biological space. This is likely to be most useful for the design of libraries specifically targeted to phenotypic screens. In this context, increasing complexity and similarity to NPs were shown to be beneficial for increasing hit rates.¹⁰¹

Phenotypic screening is an area that is continuing to see great developments. New assays that truly mimic a disease state are still challenging to achieve; however, advances in the culture and use of induced pluripotent stem cells and organoids are very promising steps in the right direction.^{102,103} Whole-organism screens in zebrafish and *Drosophila* are also gaining in traction and will add to the repertoire of options available for assessing a compound's bioactivity. Recent advances in whole-organism phenotypic screening in neurology have led to the development of the SmartCube assay system, which monitors behavioral changes in mice subjected to test compounds by using automated video acquisition and analysis.¹⁰⁴ After decades in which scientists sought to miniaturize assay systems and move away from whole-animal screening, these screening formats are beginning to increase in popularity and have seen significant technological advances.

An area that remains challenging in phenotype-driven chemical biology is the identification of a hit's target proteins. Although a range of different techniques for this purpose have been described, the best results are often obtained through the simultaneous application of multiple strategies. This is because no single unifying strategy that can successfully identify a range of target types exists. MS-based proteomic strategies are successful for soluble proteins; however, membrane proteins are

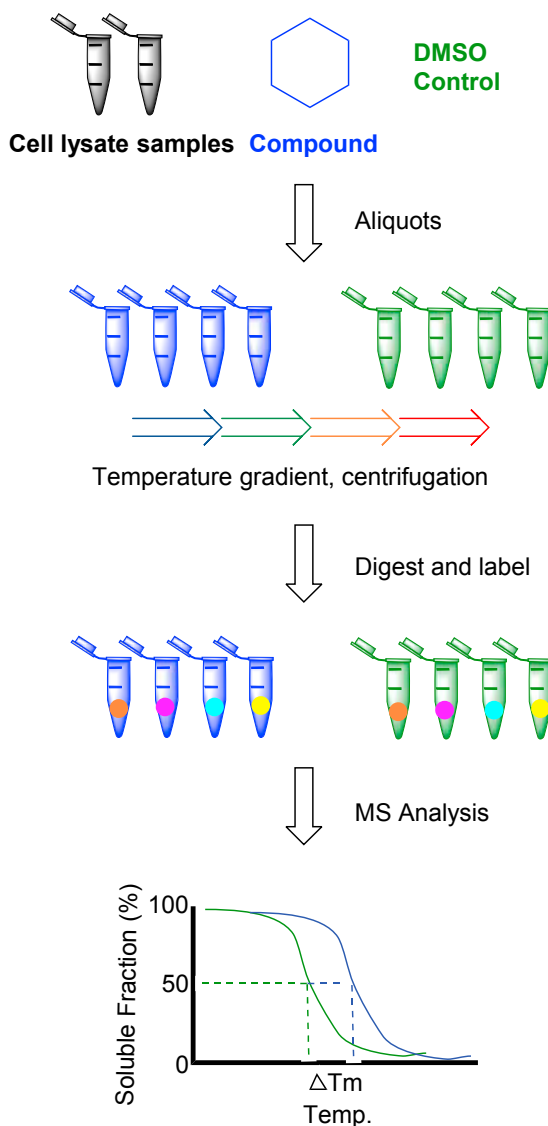


Figure 19. Thermal Proteome Profiling

Adapted from Laraia and Waldmann.⁵⁴

traditionally difficult and are best identified with covalent probes that allow more stringent washes in the subsequent purification stages. Hit prioritization remains a significant challenge because all of the described techniques typically yield multiple potential targets, which can be extremely challenging to (de)validate, depending on target type. It is also important to note that nucleic acids and lipids can also be targets for small molecules, and methods of identifying and validating them are still in their infancy. This also raises the question of which chemical matter is required for targeting these entities and whether the current screening collections contain such compounds or not in an academic and commercial setting.

In summary, despite the challenges outlined, new methods for the synthesis and identification of biologically active small molecules continue to be developed. Advances in synthesis, automation, cheminformatics, screening, bioinformatics,

and target identification mean that the field will rise in importance and relevance and will continue to see growth and success for years to come.

ACKNOWLEDGMENTS

Research in the Waldmann lab is supported by the Max-Planck-Gesellschaft. L.L. is grateful to the Alexander von Humboldt Foundation for a fellowship. L.R. is grateful to the Boehringer Ingelheim Fonds for a fellowship.

AUTHOR CONTRIBUTIONS

L.L. and H.W. conceived and coordinated the review. L.L., L.R., and H.W. wrote the manuscript. L.L. and L.R. created the figures. All authors read, revised, and approved the manuscript.

REFERENCES AND NOTES

1. Petrone, P.M., Simms, B., Nigsch, F., Lounkine, E., Kutchukian, P., Cornett, A., Deng, Z., Davies, J.W., Jenkins, J.L., and Glick, M. (2012). Rethinking molecular similarity: comparing compounds on the basis of biological activity. *ACS Chem. Biol.* 7, 1399–1409.
2. Wassermann, A.M., Lounkine, E., Urban, L., Whitebread, S., Chen, S., Hughes, K., Guo, H., Kutlina, E., Fekete, A., Klumpp, M., and Glick, M. (2014). A screening pattern recognition method finds new and divergent targets for drugs and natural products. *ACS Chem. Biol.* 9, 1622–1631.
3. Newman, D.J., and Cragg, G.M. (2012). Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* 75, 311–335.
4. Newman, D.J., and Cragg, G.M. (2016). Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* 79, 629–661.
5. Eder, J., Sedrani, R., and Wiesmann, C. (2014). The discovery of first-in-class drugs: origins and evolution. *Nat. Rev. Drug Discov.* 13, 577–587.
6. Wetzel, S., Bon, R.S., Kumar, K., and Waldmann, H. (2011). Biology-oriented synthesis. *Angew. Chem. Int. Ed.* 50, 10800–10826.
7. Dančik, V., Seiler, K.P., Young, D.W., Schreiber, S.L., and Clemons, P.A. (2010). Distinct biological network properties between the targets of natural products and disease genes. *J. Am. Chem. Soc.* 132, 9259–9261.
8. Rodrigues, T., Reker, D., Schneider, P., and Schneider, G. (2016). Counting on natural products for drug design. *Nat. Chem.* 8, 531–541.
9. van Hattum, H., and Waldmann, H. (2014). Biology-oriented synthesis: harnessing the power of evolution. *J. Am. Chem. Soc.* 136, 11853–11859.
10. Yu, M.J., Zheng, W., and Seletsky, B.M. (2013). From micrograms to grams: scale-up synthesis of eribulin mesylate. *Nat. Prod. Rep.* 30, 1158–1164.
11. Wilson, R.M., and Danishefsky, S.J. (2006). Small molecule natural products in the discovery of therapeutic agents: the synthesis connection. *J. Org. Chem.* 71, 8329–8351.
12. Lovering, F., Bikker, J., and Humblet, C. (2009). Escape from flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.* 52, 6752–6756.
13. Stratton, C.F., Newman, D.J., and Tan, D.S. (2015). Cheminformatic comparison of approved drugs from natural product versus synthetic origins. *Bioorg. Med. Chem. Lett.* 25, 4802–4807.
14. Newman, D.J., and Cragg, G.M. (2007). Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* 3, 461–477.
15. Morrison, K.C., and Hergenrother, P.J. (2014). Natural products as starting points for the synthesis of complex and diverse compounds. *Nat. Prod. Rep.* 31, 6–14.
16. Rafferty, R.J., Hicklin, R.W., Maloof, K.A., and Hergenrother, P.J. (2014). Synthesis of complex and diverse compounds through ring distortion of abietic acid. *Angew. Chem. Int. Ed.* 53, 220–224.
17. Garcia, A., Drown, B.S., and Hergenrother, P.J. (2016). Access to a structurally complex compound collection via ring distortion of the alkaloid sinomenine. *Org. Lett.* 18, 4852–4855.
18. Wang, Q., and Li, X.K. (2011). Immunosuppressive and anti-inflammatory activities of sinomenine. *Int. Immunopharmacol.* 11, 373–376.
19. Amirkia, V., and Heinrich, M. (2014). Alkaloids as drug leads – a predictive structural and biodiversity-based analysis. *Phytochem. Lett.* 10 (Suppl. C), xlviii–liii.
20. Huigens, R.W., III, Morrison, K.C., Hicklin, R.W., Flood, T.A., Jr., Richter, M.F., and Hergenrother, P.J. (2013). A ring-distortion strategy to construct stereochemically complex and structurally diverse compounds from natural products. *Nat. Chem.* 5, 195–202.
21. Waldmann, H., Khedkar, V., Dücker, H., Schürmann, M., Oppel, I.M., and Kumar, K. (2008). Asymmetric synthesis of natural product inspired tricyclic benzopyrones by an organocatalyzed annulation reaction. *Angew. Chem. Int. Ed.* 120, 6975–6978.
22. Laraia, L., Ohsawa, K., Konstantinidis, G., Robke, L., Wu, Y.W., Kumar, K., and Waldmann, H. (2017). Discovery of novel cinchona-alkaloid-inspired oxazatwistane autophagy inhibitors. *Angew. Chem. Int. Ed.* 56, 2145–2150.
23. Basu, S., Ellinger, B., Rizzo, S., Deraeve, C., Schürmann, M., Preut, H., Arndt, H.D., and Waldmann, H. (2011). Biology-oriented synthesis of a natural-product inspired oxepane collection yields a small-molecule activator of the Wnt-pathway. *Proc. Natl. Acad. Sci. USA* 108, 6805–6810.
24. Koch, M.A., Schuffenhauer, A., Scheck, M., Wetzel, S., Casaulta, M., Odermatt, A., Ertl, P., and Waldmann, H. (2005). Charting biologically relevant chemical space: a structural classification of natural products (SCONP). *Proc. Natl. Acad. Sci. USA* 102, 17272–17277.
25. Wetzel, S., Klein, K., Renner, S., Rauh, D., Oprea, T.I., Mutzel, P., and Waldmann, H. (2009). Interactive exploration of chemical space with Scaffold Hunter. *Nat. Chem. Biol.* 5, 581–583.
26. Svenda, J., Sheremet, M., Kremer, L., Maier, L., Bauer, J.O., Strohmann, C., Ziegler, S., Kumar, K., and Waldmann, H. (2015). Biology-oriented synthesis of a withanolide-inspired compound collection reveals novel modulators of hedgehog signaling. *Angew. Chem. Int. Ed.* 54, 5596–5602.
27. Zhao, J., Nakamura, N., Hattori, M., Kuboyama, T., Tohda, C., and Komatsu, K. (2002). Withanolide derivatives from the roots of *Withania somnifera* and their neurite outgrowth activities. *Chem. Pharm. Bull.* 50, 760–765.
28. Chang, H.C., Chang, F.R., Wang, Y.C., Pan, M.R., Hung, W.C., and Wu, Y.C. (2007). A bioactive withanolide Tubocapsanolide A inhibits proliferation of human lung cancer cells via repressing Skp2 expression. *Mol. Cancer Ther.* 6, 1572–1578.
29. Kuboyama, T., Tohda, C., and Komatsu, K. (2005). Neurite regeneration and synaptic reconstruction induced by withanolide A. *Br. J. Pharmacol.* 144, 961–971.
30. Kuboyama, T., Tohda, C., and Komatsu, K. (2006). Withanoid IV and its active

- metabolite, sominone, attenuate A β (25-35)-induced neurodegeneration. *Eur. J. Neurosci.* 23, 1417–1426.
31. Waldmann, H., Kapoor, S., Schröder, P., Kumar, K., Ziegler, S., and Waldmann, H. (2017). Natural product withanolide-inspired small molecules as potent inhibitors of Wnt signaling. *ChemBioChem* 18, 1797–1806.
32. Tietze, L.F., Bell, H.P., and Chandrasekhar, S. (2003). Natural product hybrids as new leads for drug discovery. *Angew. Chem. Int. Ed.* 42, 3996–4028.
33. Over, B., Wetzel, S., Grütter, C., Nakai, Y., Renner, S., Rauh, D., and Waldmann, H. (2013). Natural-product-derived fragments for fragment-based ligand discovery. *Nat. Chem.* 5, 21–28.
34. Gryniewicz, G.G.M. (2008). Tropane alkaloids as medicinally useful natural products and their synthetic derivatives as new drugs. *Pharmacol. Rep.* 60, 439–463.
35. O'Hagan, D. (2000). Pyrrole, pyrrolidine, pyridine, piperidine and tropane alkaloids (1998 to 1999). *Nat. Prod. Rep.* 17, 435–446.
36. Narayan, R., Bauer, J.O., Strohmman, C., Antonchick, A.P., and Waldmann, H. (2013). Catalytic enantioselective synthesis of functionalized tropanes reveals novel inhibitors of hedgehog signaling. *Angew. Chem. Int. Ed.* 52, 12892–12896.
37. Xu, H., Golz, C., Strohmman, C., Antonchick, A.P., and Waldmann, H. (2016). Enantiodivergent combination of natural product scaffolds enabled by catalytic enantioselective cycloaddition. *Angew. Chem. Int. Ed.* 55, 7761–7765.
38. Schreiber, S.L. (2009). Organic chemistry: molecular diversity by design. *Nature* 457, 143–154.
39. O'Connell, K.M., Beckmann, H.S., Laraia, L., Horsley, H.T., Bender, A., Venkitaraman, A.R., and Spring, D.R. (2012). A two-directional strategy for the diversity-oriented synthesis of macrocyclic scaffolds. *Org. Biomol. Chem.* 10, 7545–7551.
40. Galloway, W.R., Isidro-Llobet, A., and Spring, D.R. (2010). Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules. *Nat. Commun.* 1, 80.
41. Garcia-Castro, M., Kremer, L., Reinkemeier, C.D., Unkelbach, C., Strohmman, C., Ziegler, S., Ostermann, C., and Kumar, K. (2015). De novo branching cascades for structural and functional diversity in small molecules. *Nat. Commun.* 6, 6516.
42. Kim, J., Kim, H., and Park, S.B. (2014). Privileged structures: efficient chemical "Navigators" toward unexplored biologically relevant chemical spaces. *J. Am. Chem. Soc.* 136, 14629–14638.
43. Kim, J., Jung, J., Koo, J., Cho, W., Lee, W.S., Kim, C., Park, W., and Park, S.B. (2016). Diversity-oriented synthetic strategy for developing a chemical modulator of protein-protein interaction. *Nat. Commun.* 7, 13196.
44. Yaziji, V., Rodríguez, D., Gutiérrez-de-Terán, H., Coelho, A., Caamaño, O., García-Mera, X., Brea, J., Loza, M.I., Cadavid, M.I., and Sotelo, E. (2011). Pyrimidine derivatives as potent and selective A3 adenosine receptor antagonists. *J. Med. Chem.* 54, 457–471.
45. Robke, L., Laraia, L., Carnero Corrales, M.A., Konstantinidis, G., Muroi, M., Richters, A., Winzker, M., Engbring, T., Tomassi, S., Watanabe, N., et al. (2017). Phenotypic identification of a novel autophagy inhibitor chemotype targeting lipid kinase VPS34. *Angew. Chem. Int. Ed.* 56, 8153–8157.
46. Vitaku, E., Smith, D.T., and Njardarson, J.T. (2014). Analysis of the structural diversity, substitution patterns, and frequency of nitrogen heterocycles among U.S. FDA approved pharmaceuticals. *J. Med. Chem.* 57, 10257–10274.
47. Liu, W., Khedkar, V., Baskar, B., Schürmann, M., and Kumar, K. (2011). Branching cascades: a concise synthetic strategy targeting diverse and complex molecular frameworks. *Angew. Chem. Int. Ed.* 50, 6900–6905.
48. Wang, M., Zhang, X., Zhuang, Y.X., Xu, Y.H., and Loh, T.P. (2015). Pd-catalyzed intramolecular C–N bond cleavage, 1,4-migration, sp³ C–H activation, and Heck reaction: four controllable diverse pathways depending on the judicious choice of the base and ligand. *J. Am. Chem. Soc.* 137, 1341–1347.
49. Morton, D., Leach, S., Cordier, C., Warriner, S., and Nelson, A. (2009). Synthesis of natural-product-like molecules with over eighty distinct scaffolds. *Angew. Chem. Int. Ed.* 48, 104–109.
50. Lee, Y.C., Patil, S., Golz, C., Strohmman, C., Ziegler, S., Kumar, K., and Waldmann, H. (2017). A ligand-directed divergent catalytic approach to establish structural and functional scaffold diversity. *Nat. Commun.* 8, 14043.
51. Alza, E., Laraia, L., Ibbeson, B.M., Collins, S., Galloway, W.R.J.D., Stokes, J.E., Venkitaraman, A.R., and Spring, D.R. (2015). Synthesis of a novel polycyclic ring scaffold with antimetabolic properties via a selective domino Heck-Suzuki reaction. *Chem. Sci.* 6, 390–396.
52. Ibbeson, B.M., Laraia, L., Alza, E., O'Connor, C.J., Tan, Y.S., Davies, H.M., McKenzie, G., Venkitaraman, A.R., and Spring, D.R. (2014). Diversity-oriented synthesis as a tool for identifying new modulators of mitosis. *Nat. Commun.* 5, 3155.
53. Moffat, J.G., Vincent, F., Lee, J.A., Eder, J., and Prunotto, M. (2017). Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat. Rev. Drug Discov.* 16, 531–543.
54. Laraia, L., and Waldmann, H. (2017). Natural product inspired compound collections: evolutionary principle, chemical synthesis, phenotypic screening, and target identification. *Drug Discov. Today Technol.* 23, 75–82.
55. Mayer, T.U., Kapoor, T.M., Haggarty, S.J., King, R.W., Schreiber, S.L., and Mitchison, T.J. (1999). Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* 286, 971.
56. Galluzzi, L., Bravo-San Pedro, J.M., Levine, B., Green, D.R., and Kroemer, G. (2017). Pharmacological modulation of autophagy: therapeutic potential and persisting obstacles. *Nat. Rev. Drug Discov.* 16, 487–511.
57. Levy, J.M.M., Towers, C.G., and Thorburn, A. (2017). Targeting autophagy in cancer. *Nat. Rev. Cancer* 17, 528–542.
58. Xu, H., Laraia, L., Schneider, L., Louven, K., Strohmman, C., Antonchick, A.P., and Waldmann, H. (2017). Highly enantioselective catalytic vinylogous propargylation of coumarins yields a class of autophagy inhibitors. *Angew. Chem. Int. Ed.* 56, 11232–11236.
59. Liu, J., Xia, H., Kim, M., Xu, L., Li, Y., Zhang, L., Cai, Y., Norberg, H.V., Zhang, T., Furuya, T., et al. (2011). Beclin1 controls the levels of p53 by regulating the deubiquitination activity of USP10 and USP13. *Cell* 147, 223–234.
60. Gustafsdottir, S.M., Ljosa, V., Sokolnicki, K.L., Anthony Wilson, J., Walpita, D., Kemp, M.M., Petri Seiler, K., Carrel, H.A., Golub, T.R., Schreiber, S.L., et al. (2013). Multiplex cytological profiling assay to measure diverse cellular states. *PLoS One* 8, e80999.
61. Reisen, F., Sauty de Chalon, A., Pfeifer, M., Zhang, X., Gabriel, D., and Selzer, P. (2015). Linking phenotypes and modes of action through high-content screen fingerprints. *Assay Drug Dev. Technol.* 13, 415–427.
62. Bray, M.-A., Singh, S., Han, H., Davis, C.T., Borgeson, B., Hartland, C., Kost-Alimova, M., Gustafsdottir, S.M., Gibson, C.C., and Carpenter, A.E. (2016). Cell painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat. Protoc.* 11, 1757–1774.
63. Caicedo, J.C., Cooper, S., Heigwer, F., Warchal, S., Qiu, P., Molnar, C., Vasilevich, A.S., Barry, J.D., Bansal, H.S., Kraus, O., et al. (2017). Data-analysis strategies for image-based cell profiling. *Nat. Methods* 14, 849–863.
64. Wawer, M.J., Li, K., Gustafsdottir, S.M., Ljosa, V., Bodycombe, N.E., Marton, M.A., Sokolnicki, K.L., Bray, M.A., Kemp, M.M., Winchester, E., et al. (2014). Toward performance-diverse small-molecule libraries for cell-based phenotypic screening using multiplexed high-dimensional profiling. *Proc. Natl. Acad. Sci. USA* 111, 10911–10916.
65. Futamura, Y., Kawatani, M., Kazami, S., Tanaka, K., Muroi, M., Shimizu, T., Tomita, K., Watanabe, N., and Osada, H. (2012). Morphobase, an encyclopedic cell morphology database, and its use for drug target identification. *Chem. Biol.* 19, 1620–1630.
66. Rimkus, T.K., Carpenter, R.L., Qasem, S., Chan, M., and Lo, H.W. (2016). Targeting the sonic hedgehog signaling pathway: review of smoothened and GLI inhibitors. *Cancers (Basel)* 8, <https://doi.org/10.3390/cancers8020022>.
67. Ziegler, S., Pries, V., Hedberg, C., and Waldmann, H. (2013). Target identification for small bioactive molecules: finding the needle in the haystack. *Angew. Chem. Int. Ed.* 52, 2744–2792.

68. Keiser, M.J., Roth, B.L., Armbruster, B.N., Ernsberger, P., Irwin, J.J., and Shoichet, B.K. (2007). Relating protein pharmacology by ligand chemistry. *Nat. Biotech.* 25, 197–206.
69. Gfeller, D., Grosdidier, A., Wirth, M., Daina, A., Michielin, O., and Zoete, V. (2014). SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucleic Acids Res.* 42, W32–W38.
70. Nickel, J., Gohlke, B.O., Erethman, J., Banerjee, P., Rong, W.W., Goede, A., Dunkel, M., and Preissner, R. (2014). SuperPred: update on drug classification and target prediction. *Nucleic Acids Res.* 42, W26–W31.
71. Kringelum, J., Kjaerulf, S.K., Brunak, S., Lund, O., Oprea, T.I., and Taboureau, O. (2016). ChemProt-3.0: a global chemical biology diseases mapping. *Database* 2016, <https://doi.org/10.1093/database/bav123>.
72. Reker, D., Rodrigues, T., Schneider, P., and Schneider, G. (2014). Identifying the macromolecular targets of de novo-designed chemical entities through self-organizing map consensus. *Proc. Natl. Acad. Sci. USA* 111, 4067–4072.
73. Kremer, L., Schultz-Fademrecht, C., Baumann, M., Habenberger, P., Choidas, A., Klebl, B., Kordes, S., Schöler, H.R., Sterneckert, J., Ziegler, S., et al. (2017). Discovery of a novel hedgehog signaling pathway inhibitor by cell-based compound discovery and target prediction. *Angew. Chem. Int. Ed.* 56, 13021–13025.
74. Wang, L., Ma, C., Wipf, P., Liu, H., Su, W., and Xie, X.Q. (2013). TargetHunter: an in silico target identification tool for predicting therapeutic potential of small organic molecules based on chemogenomic database. *AAPS J.* 15, 395–406.
75. Lo, Y.C., Senese, S., Li, C.M., Hu, Q., Huang, Y., Damoiseaux, R., and Torres, J.Z. (2015). Large-scale chemical similarity networks for target profiling of compounds identified in cell-based chemical screens. *PLoS Comp. Biol.* 11, e1004153.
76. Ursu, A., Illich, D.J., Takemoto, Y., Porfetye, A.T., Zhang, M., Brockmeyer, A., Janning, P., Watanabe, N., Osada, H., Vetter, I.R., et al. (2016). Epiblastin induces reprogramming of epiblast stem cells into embryonic stem cells by inhibition of Casein kinase 1. *Cell Chem. Biol.* 23, 494–507.
77. Schröder, P., Förster, T., Kleine, S., Becker, C., Richters, A., Ziegler, S., Rauh, D., Kumar, K., and Waldmann, H. (2015). Neuritogenic Militarone-inspired 4-Hydroxypyridones target the stress pathway kinase MAP4K4. *Angew. Chem. Int. Ed.* 54, 12398–12403.
78. Kroeze, W.K., Sassano, M.F., Huang, X.P., Lansu, K., McCovry, J.D., Giguère, P.M., Sciaky, N., and Roth, B.L. (2015). PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat. Struct. Mol. Biol.* 22, 362–369.
79. Barnea, G., Strapps, W., Herrada, G., Berman, Y., Ong, J., Kloss, B., Axel, R., and Lee, K.J. (2008). The genetic design of signaling cascades to record receptor activation. *Proc. Natl. Acad. Sci. USA* 105, 64–69.
80. Ong, S.-E., and Mann, M. (2007). A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* 1, 2650–2660.
81. Kapoor, S., Waldmann, H., and Ziegler, S. (2016). Novel approaches to map small molecule–target interactions. *Bioorg. Med. Chem.* 24, 3232–3245.
82. Schenone, M., Dančik, V., Wagner, B.K., and Clemons, P.A. (2013). Target identification and mechanism of action in chemical biology and drug discovery. *Nat. Chem. Biol.* 9, 232–240.
83. Smith, E., and Collins, I. (2015). Photoaffinity labeling in target- and binding-site identification. *Future Med. Chem.* 7, 159–183.
84. Park, J., Oh, S., and Park, S.B. (2012). Discovery and target identification of an antiproliferative agent in live cells using fluorescence difference in two-dimensional gel electrophoresis. *Angew. Chem. Int. Ed.* 51, 5447–5451.
85. Parker, C.G., Galmozzi, A., Wang, Y., Correia, B.E., Sasaki, K., Joslyn, C.M., Kim, A.S., Cavallaro, C.L., Lawrence, R.M., Johnson, S.R., et al. (2017). Ligand and target discovery by fragment-based screening in human cells. *Cell* 168, 527–541.
86. Savitski, M.M., Reinhard, F.B., Franken, H., Werner, T., Savitski, M.F., Eberhard, D., Martinez Molina, D., Jafari, R., Dovega, R.B., Klaeger, S., et al. (2014). Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* 346, <https://doi.org/10.1126/science.1255784>.
87. Franken, H., Mathieson, T., Childs, D., Sweetman, G.M., Werner, T., Tögel, I., Doce, C., Gade, S., Bantscheff, M., Drewes, G., et al. (2015). Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat. Protoc.* 10, 1567.
88. Werner, T., Sweetman, G., Savitski, M.F., Mathieson, T., Bantscheff, M., and Savitski, M.M. (2014). Ion coalescence of neutron encoded TMT 10-Plex reporter ions. *Anal. Chem.* 86, 3594–3601.
89. Park, H., Ha, J., Koo, J.Y., Park, J., and Park, S.B. (2017). Label-free target identification using in-gel fluorescence difference via thermal stability shift. *Chem. Sci.* 8, 1127–1133.
90. Muroi, M., Futamura, Y., and Osada, H. (2016). Integrated profiling methods for identifying the targets of bioactive compounds: MorphoBase and ChemProteoBase. *Nat. Prod. Rep.* 33, 621–625.
91. Futamura, Y., Kawatani, M., Muroi, M., Aono, H., Nogawa, T., and Osada, H. (2013). Identification of a molecular target of a novel fungal metabolite, pyrrolizilactone, by phenotypic profiling systems. *ChemBioChem* 14, 2456–2463.
92. Smith, A.M., Ammar, R., Nislow, C., and Giaever, G. (2010). A survey of yeast genomic assays for drug and target discovery. *Pharm. Ther.* 127, 156–164.
93. Nijman, S.M.B. (2015). Functional genomics to uncover drug mechanism of action. *Nat. Chem. Biol.* 11, 942–948.
94. Andersson, B., Collins, V.P., Kurzrock, R., Larkin, D.W., Childs, C., Ost, A., Cork, A., Trujillo, J.M., Freireich, E.J., Siciliano, M.J., et al. (1996). KBM-7, a human myeloid leukemia cell line with double Philadelphia chromosomes lacking normal c-ABL and BCR transcripts. *Leukemia* 9, 2100–2108.
95. Smurnyy, Y., Cai, M., Wu, H., McWhinnie, E., Tallarico, J.A., Yang, Y., and Feng, Y. (2014). DNA sequencing and CRISPR-Cas9 gene editing for target validation in mammalian cells. *Nat. Chem. Biol.* 10, 623–625.
96. Dixon, S.J., Winter, G.E., Musavi, L.S., Lee, E.D., Snijder, B., Rebsamen, M., Superti-Furga, G., and Stockwell, B.R. (2015). Human haploid cell genetics reveals roles for lipid metabolism genes in nonapoptotic cell death. *ACS Chem. Biol.* 10, 1604–1609.
97. Winter, G.E., Radic, B., Mayor-Ruiz, C., Blomen, V.A., Trefzer, C., Kandasamy, R.K., Huber, K.V.M., Gridling, M., Chen, D., Klampfl, T., et al. (2014). The solute carrier SLC35F2 enables YM155-mediated DNA damage toxicity. *Nat. Chem. Biol.* 10, 768–773.
98. Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483, 603–607.
99. Carson, C., Raman, P., Tullai, J., Xu, L., Henault, M., Thomas, E., Yeola, S., Lao, J., McPate, M., Verkuy, J.M., et al. (2015). Englerin agonizes the TRPC4/C5 cation channels to inhibit tumor cell line proliferation. *PLoS One* 10, e0127498.
100. Akbulut, Y., Gaunt, H.J., Muraki, K., Ludlow, M.J., Amer, M.S., Bruns, A., Vasudev, N.S., Radtke, L., Willot, M., Hahn, S., et al. (2015). (–)-Englerin a is a potent and selective activator of TRPC4 and TRPC5 calcium channels. *Angew. Chem. Int. Ed.* 54, 3787–3791.
101. Spear, K.L., and Brown, S.P. (2017). The evolution of library design: crafting smart compound collections for phenotypic screens. *Drug Discov. Today Technol.* 23 (Suppl. C), 61–67.
102. Ursu, A., Schöler, H.R., and Waldmann, H. (2017). Small-molecule phenotypic screening with stem cells. *Nat. Chem. Biol.* 13, 560–563.
103. Fatehullah, A., Tan, S.H., and Barker, N. (2016). Organoids as an in vitro model of human development and disease. *Nat. Cell Biol.* 18, 246–254.
104. Alexandrov, V., Brunner, D., Hanania, T., and Leahy, E. (2015). High-throughput analysis of behavior for drug discovery. *Eur. J. Pharmacol.* 750 (Suppl. C), 82–89.